# Essays on Panel Cointegration Testing

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum politicarum
(Dr. rer. pol.)
im Fach Wirtschaftswissenschaften

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
Humboldt-Universität zu Berlin

von
Frau M.Sc. Deniz Dilan Karaman Örsal
geboren am 18.10.1977 in Ankara, Türkei

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Christoph Markschies

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Oliver Günther Ph.D.

Gutachter:

1. Priv.-Doz. Dr. Bernd Droge
2. Prof. Dr. Nikolaus Hautsch

eingereicht am: 28. Juli 2008
Tag der mündlichen Prüfung: 4. Februar 2009

**Abstract**

This thesis is composed of four essays which contribute to the literature in panel cointegration methodology. The first essay compares the finite sample properties of the four residual-based panel cointegration tests of Pedroni (1995, 1999) and the likelihood-based panel cointegration test of Larsson et al. (2001). The simulation results indicate that the panel-$t$ test statistic of Pedroni (1995, 1999) has the best finite sample properties among the five panel cointegration test statistics evaluated. The second essay presents a corrected version of the proof of Larsson et al. (2001) related to the finiteness of the moments of the asymptotic trace statistic. The proof is corrected for the case, in which the difference between the number of variables and the number of existing cointegrating relations is one. The third essay proposes a new likelihood-based panel cointegration test in the presence of a linear time trend in the data generating process. This new test is an extension of the likelihood ratio test of Saikkonen and Lütkepohl (2000a) for trend-adjusted data to the panel data framework, and is called the panel SL test. The idea is first to take the average of the individual LR (trace) statistics over the cross-sections and then to standardize the test statistic with the appropriate asymptotic moments. Under the null hypothesis, the panel SL test statistic is standard normally distributed as the number of time periods ($T$) and the number of cross-sections ($N$) tend to infinity sequentially. By means of a Monte Carlo study the finite sample properties of the test are investigated. The new test presents reasonable size with the increase in $T$ and $N$, and has high power in small samples. The last essay of the thesis analyzes the long-run money demand relation among OECD countries by panel unit root and cointegration testing techniques. The panel SL cointegration test and the tests of Pedroni (1999) are used to detect the existence of a stationary long-run money demand relation. Moreover, the money demand function is estimated with the panel dynamic ordinary least squares method of Mark and Sul (2003).

**Keywords:**
Panel Cointegration, Monte Carlo Study, Fisher Hypothesis, Asymptotic Moments, Trace Statistic, Uniform Integrability, Uniform Boundedness, Linear Trend, Money Demand

## Zusammenfassung

Diese Dissertation beinhaltet vier Aufsätze, die zur Literatur der Panelkointegrationsmethodik beitragen. Der erste Aufsatz vergleicht die Eigenschaften der vier Residuen-basierten Panelkointegrationstests von Pedroni (1995, 1999) mit dem Likelihood-basierten Panelkointegrationstest von Larsson et al. (2001) in endlichen Stichproben. Die Simulationsergebnisse zeigen, dass unter den fünf untersuchten Panelkointegrationsteststatistiken die Panel-$t$ Teststatistik von Pedroni (1995, 1999) die besten Eigenschaften in endlichen Stichproben besitzt. Der zweite Aufsatz präsentiert eine Korrektur des Beweises von Larsson et al. (2001) bezüglich der Endlichkeit der Momente der asymptotischen Trace-Statistik für den Fall, dass die Differenz zwischen der Anzahl der Variablen und der Anzahl der existierenden Kointegrationsbeziehungen eins ist. Im dritten Aufsatz wird ein neuer Likelihood-basierter Panelkointegrationstest vorgestellt, der die Existenz eines linearen Trends in dem datengenerierenden Prozess erlaubt. Dieser neue Test ist eine Erweiterung des Likelihood-Quotienten (LR)-Tests von Saikkonen und Lütkepohl (2000a) für trendbereinigte Daten auf die Paneldatenanalyse. Die Idee dieses im Folgenden als Panel-SL Test bezeichneten Verfahrens ist, den Mittelwert aus den individuellen LR Statistiken über die Querschnitte zu bilden und danach die Teststatistik mit Hilfe der geeigneten asymptotischen Momente zu standardisieren. Unter der Nullhypothese folgt die Panel-SL Teststatistik einer standardisierten Normalverteilung, wenn die Anzahl der Beobachtungen über die Zeit ($T$) und die Anzahl der Querschnitte ($N$) sequentiell gegen unendlich gehen. In einer Monte-Carlo-Studie werden die Eigenschaften des Panel-SL Tests in endlichen Stichproben untersucht. Der neue Test hat ein annehmbares empirisches Signifikanzniveau für wachsende $T$ und $N$ sowie eine hohe Güte in kleinen Stichproben. Der letzte Aufsatz der Dissertation analysiert die langfristige Geldnachfragefunktion in OECD Ländern mit Hilfe von Paneleinheitswurzel- und Panelkointegrationstests. Um eine mögliche Existenz einer stationären langfristigen Geldnachfragefunktion zu untersuchen, werden der Panel-SL Kointegrationstest und die Tests von Pedroni (1999) verwendet. Im Anschluss daran wird eine Paneldatenschätzung für die Geldnachfragefunktion mittels der dynamischen Kleinste-Quadrate-Methode von Mark und Sul (2003) durchgeführt.


**Schlagwörter:**
Panelkointegration, Monte-Carlo Studie, Fisher-Hypothese, Asymptotische Momente, Trace-Statistik, Gleichmäßige Integrierbarkeit, Gleichmäßige Beschränktheit, Linearer Trend, Geldnachfrage

# Acknowledgments

I would like to express my thankfulness to my supervisor PD Dr. Bernd Droge, for his comments, support and encouragement while writing this thesis. I would like to thank my second examiner Prof. Dr. Nikolaus Hautsch, for his support during my studies at the Chair of Econometrics. Furthermore, I am grateful to my previous supervisor Prof. Dr. Helmut Herwartz who awakened the interest in me to do research in the field of panel cointegration.

I owe special thanks to Jana Riedel and Ozan Karaman for reading my thesis. Without their comments and suggestions I would not be able to achieve the final version of this thesis.

I am deeply grateful to the Deutsche Forschungsgemeinschaft (DFG) and the Sonderforschungsbereich SFB 649 (Economic Risk) of the Humboldt University, Berlin, for their financial support.

I address my thanks to my dear colleagues Prof. Dr. Carsten Trenkler, Dr. Danijela Markovic, Dr. Markus Krtzig, Oliver Blaskowitz and Prof. Dr. Ralf Brggemann, who supported me with their advice and comments throughout the years that I worked with them.

I am also indebted to all my friends who encouraged me to complete this thesis. I would like to express my sincerest gratitude to my mum İnci, my dad Veli, my brother Ozan and my grandma Mkerrem for the support they provided me through my entire life. Finally, I want to thank my dear husband, Arif Suphi, without whose love, faith and encouragement, I would not have finished this thesis.

# Contents

# Chapter 1

# Introduction

In the last two decades, cointegration techniques have been widely used in empirical literature. The concept of cointegration was first introduced by Engle and Granger (1987). Cointegration defines a long-run relation between integrated variables[1]. In other words, if a linear combination of the integrated variables of order $d$, is integrated of a smaller order than $d$, then these variables are cointegrated. For example, if the components of a $K$-dimensional process $y_t$ are $I(1)$ and a linear combination of the components of $y_t$ is stationary, i.e. $\beta' y_t = (\beta_1, \ldots, \beta_K)' y_t \sim I(0)$ for $\beta \neq 0$, then the variables are cointegrated.

According to another definition of cointegration if at least one of the components of a $K$-dimensional process $y_t$ is $I(d)$, then $y_t \sim I(d)$. As a result, if $\beta' y_t$ is integrated of a smaller order than $d$, the components of the process $y_t$ are cointegrated (see Lütkepohl, 2005). In this context, to test for cointegrating relation(s) between the components of a process $y_t$, it is not necessary that all the components of $y_t$ are integrated of the same order.

It is known that applications of the conventional time series techniques to integrated variables cause inefficient results. The regression of an integrated process on an unrelated integrated process delivers a high $t$-ratio of the slope parameter, which points out a significant relation between these unrelated processes. This is due to the fact that the variance of the regression cannot be estimated consistently. Additionally, the residuals of this so-called spurious regression are nonstationary. However, if the residuals of this regression are $I(0)$, then there is a cointegrating relation between the integrated variables. Hence, cointegration techniques are often more appropriate for describing economic models than the conventional time series techniques as most of the economic variables, e.g. prices, consumption, income etc., are integrated

[1] If a univariate time series process $x_t$ has $d$ unit roots, then $x_t$ is integrated of order $d$, i.e. $x_t \sim I(d)$.

processes.

There are mainly two different types of cointegration tests. The first type of tests are residual-based (single-equation) cointegration tests and the second type of tests are maximum-likelihood-based (systems) cointegration tests. The residual-based cointegration tests are used to detect the presence of a cointegrating relation between the variables. However, the maximum-likelihood-based cointegration tests are used not only to detect the presence of cointegrating relation(s), but also to determine the number of cointegrating relations between the variables of a system. Another advantage of the maximum-likelihood-based cointegration tests over the residual-based tests is the independence of the tests to the choice of the variable used for the normalization of the cointegrating relation.

Cointegration tests are extensions of the unit root tests to the multivariate time series framework. Unit root tests are tools to determine the order of integratedness of a univariate time series. It is known that the standard Dickey-Fuller (DF) and Augmented Dickey-Fuller (ADF) unit root tests have low power against near unit root alternatives for small samples. Perron (1989, 1991), Shiller and Perron (1985) and Pierce and Snell (1995) found that the power of these tests can be increased when the time span of the data increases. They also concluded that the time span of the data but not the frequency is important for increasing the power of the tests (Pedroni, 2004).

As known, if the span of the time series increases, problems related to structural breaks and regime shifts may occur. Moreover, sometimes the data availability for a cross-section may be limited. Hence, another way to increase the power of the tests and the number of observations available may be to add data from different cross-sections e.g. countries, firms, industries etc., which leads us to the panel data unit root and cointegration tests.

The first panel unit root tests were introduced by Levin and Lin (1993), Quah (1994), Breitung and Meyer (1994) and Im et al. (1997). Their tests are extensions of the standard DF and ADF unit root tests to a panel data framework. They test the null hypothesis of a unit root process against the alternative of a stationary process, allowing different degrees of heterogeneity over cross-sections. These test statistics are mainly the standardized versions of the mean of the DF and ADF statistics, which are asymptotically standard normally distributed as the time ($T$) and the cross-section ($N$) dimensions go to infinity.

The panel unit root tests for the null hypothesis of a stationary process against the alternative of a unit root process were proposed by Nyblom and Harvey (2000) and Choi (2001). This was followed by the development of tests which allow for cross-sectional dependence (Maddala and Wu, 1999; Chang, 2002; Chang and Park, 2003; Chang, 2004) and structural breaks

(Im et al., 2005; Jönsson, 2005; Costantini and Gutierrez, 2007) to prevent size distortions due to neglecting cross-sectional dependence and structural breaks.

Similar to panel unit root tests such as those of Levin and Lin (1993) and Im et al. (2003), Pedroni (1995, 1999) and Kao (1999) based the first panel cointegration tests on DF and ADF statistics. In contrast to the standard multivariate cointegration tests, these residual-based tests do not suffer from low power in small samples, at least if the cross-section dimension is large. Moreover, Pedroni (1995, 1999) introduced also panel cointegration tests, which are extensions of the variance ratio test of Phillips and Ouliaris (1990) and the $\rho$-tests of Phillips and Perron (1988), Phillips and Ouliaris (1990). In addition to this, the first residual-based cointegration test for the null hypothesis of cointegration was proposed by McCoskey and Kao (1998), which is built on the tests of Harris and Inder (1994) and Shin (1994).

The maximum-likelihood-based panel cointegration tests are based on the following vector error correction model (VECM):

$$\begin{aligned} \Delta y_{it} &= \alpha_i \beta_i' y_{i,t-1} + \sum_{j=1}^{p_i-1} \Gamma_{ij} \Delta y_{i,t-j} + C_i d_t + \varepsilon_{it}, \\ & i = 1, \ldots, N; \quad t = 1, \ldots, T, \end{aligned} \tag{1.1}$$

in which $i$ denotes the index for cross-section, $t$ is the index for time and the error terms $\varepsilon_{it}$ are independently distributed, $\varepsilon_{it} \sim N_K(0, \Omega_i)$. The components of the $K$-dimensional process $y_{it}$ are at most $I(1)$ and cointegrated with cointegrating rank $r_i$ $(0 \leq r_i \leq K)$. The unknown $(K \times r_i)$ matrices $\alpha_i$ and $\beta_i$ are the loading and the cointegrating matrices, respectively, with full column rank. The lag order of the vector error correction (VEC) process, i.e. $p_i - 1$, is either cross-sectionally variant or restricted to be constant, i.e. $p_i = p$ for all $i$. The unknown coefficient matrices $\Gamma_{ij}$, $i = 1, \ldots, N$; $j = 1, \ldots, p_i - 1$, denote the short-run dynamics of the process, whereas $\Pi_i = \alpha_i \beta_i'$, $i = 1, \ldots, N$, represent the long-run dynamics of the system. Finally, the vector $d_t$ contains the deterministic terms and $C_i$ is the unknown parameter vector of the deterministic terms.

The first maximum-likelihood-based panel cointegration test was suggested by Larsson et al. (2001), and is an extension of the trace statistic of Johansen (1995) to heterogeneous panel data. Under the null hypothesis Larsson et al. (2001) test $\text{rank}(\Pi_i) = r_i \leq r$ for all $i$, assuming that there is no cross-sectional dependence, i.e $E(\varepsilon_{it} \varepsilon_{jt}') = 0$ for $i \neq j$. Their panel test statistic is a standardized version of the average of the individual trace statistics, where the first two moments of the asymptotic trace statistic is used to standardize. Hence, they called their panel cointegration statistic standardized LR-bar statistic, where LR denotes likelihood ratio. However, the test of Larsson et al. (2001) does not allow for deterministic terms in the cointegrating relation, i.e. $d_t = 0$. Under certain conditions the standardized LR-bar statistic is asymptotically $N(0,1)$ distributed as $T$ and $N$ go to infinity simultaneously, in such a way that $\sqrt{N}/T \to 0$. To establish this asymptotic distribution one of the important conditions is the existence and finiteness of the first two moments of the asymptotic trace statistic.

A second maximum-likelihood-based panel cointegration test was suggested by Groen and Kleibergen (2003), which is also an LR statistic for common cointegrating rank (i.e. $r_i = r$ for all $i$), and common cointegrating vector (i.e. $\beta_i = \beta$ for all $i$). This test allows for cross-sectional dependence (i.e. $E(\varepsilon_{it}\varepsilon'_{jt}) \neq 0$ for $i \neq j$) and deterministic terms. The estimation procedure is based on the generalized method of moments (GMM) estimator of Hansen (1982). Following a similar procedure as in Larsson et al. (2001), the panel cointegration statistic is standardized with asymptotic moments provided in Larsson et al. (2001). Asymptotically the standardized panel test is $N(0,1)$ distributed as $T$ goes to infinity, when $N$ is fixed.

Another systems panel cointegration testing procedure was introduced by Breitung (2005). He proposed LR, Lagrange multiplier (LM) and Wald tests, which are based on the procedure of Saikkonen (1999). The cointegrated vector autoregressive (VAR) model does not allow for short-run dynamics, i.e $\Gamma_{ij} = 0$ for $i = 1, \ldots, N$; $j = 1, \ldots p_i - 1$. Moreover, the cointegrating vectors are assumed to be common over cross-sections. The standardized versions of the test statistics are asymptotically standard normally distributed as $T$ and $N$ go to infinity sequentially. To standardize the test statistics Breitung (2005) also used the asymptotic moments from Larsson et al. (2001) because the test statistics of Breitung (2005) without any short-run dynamics and deterministic terms has the same asymptotic distribution as the test statistic of Larsson et al. (2001). In addition to this, Breitung (2005) assumed that the tests could be extended to the cases with short-run dynamics and/or deterministic terms, but did not deliver proof on the asymptotic distributions and the finiteness of the asymptotic moments.

## 1.1 Objective

The main aim of this thesis is to contribute to the literature on panel cointegration testing. Besides an extensive simulation study on the properties of some panel cointegration tests, the thesis also comprises the correction of the proof on the existence and finiteness of the moments of the asymptotic trace statistic, and offers a new maximum-likelihood-based panel cointegration test which allows a linear time trend.

The first aim of this study is to compare the finite sample size and size-adjusted power properties of the most frequently used panel cointegration tests in literature. By means of a Monte Carlo study, the panel-$\rho$, the group-$\rho$, the parametric panel-$t$, the parametric group-$t$ tests of Pedroni (1995, 1999) and the standardized LR-bar test of Larsson et al. (2001) are compared. The results from this Monte Carlo study has been published in Karaman Örsal (2008). Additionally, a long-run relationship between the nominal interest rate and the inflation rate for a panel dataset consisting of OECD countries comprising the period from June 1989 to December 20002 is tested.

The second aim of this study is to correct the proof of Lemma 1 in Larsson et al. (2001) related to the existence and finiteness of the moments of the asymptotic trace statistic. First, the reasons why the proof is incorrect are provided. Then, the proof of Larsson et al. (2001) is corrected for the case, in which the difference between the number of variables ($K$) and the number of existing cointegrating relations ($r$) is one, i.e. $K-r=1$. The proof, which is based on the Cauchy-Schwarz inequality and a sufficient condition for the uniform integrability of the random variables shows that the moments of the asymptotic trace statistic for $K-r=1$ exist and are finite. Moreover, the moments of the trace statistic converge to the moments of the asymptotic trace statistic as the number of observations (i.e. $T$) approaches infinity. Following a similar procedure the finiteness of the asymptotic moments of the test statistics for some panel unit root tests can also be proved.

The third aim of this study is to propose a new maximum-likelihood-based panel cointegration test, which allows for short-run dynamics and deterministic terms in the heterogeneous VAR model. This new test is an extension of the generalized least squares (GLS)-based LR test[2] of Saikkonen and Lütkepohl (2000a) to panel data framework, and is called the panel SL test. The idea is first to take the average of the individual trace statistics over cross-sections and then to standardize the test statistic with the appropriate asymptotic moments. This standardized panel cointegration statistic has a limiting normal distribution as $T$ and $N$ tend to infinity sequentially. In addition to the approximation based on asymptotic moments, a second approximation approach involving the moments from a VAR(1) process is introduced. By a Monte Carlo study the finite sample size and size-adjusted power properties of the tests are analyzed. The properties of the panel SL test is also compared with the properties of the Larsson et al. (2001) test which

[2]The principle of the GLS-based LR test is to subtract the GLS estimates of the deterministic terms from the original data and applying the cointegration test on the trend-adjusted data.

allows for deterministic terms[3]. Furthermore, the proof on the existence and finiteness of the moments of the asymptotic GLS-based LR statistic is also demonstrated, which is built on the corrected proof of Larsson et al. (2001).

The last contribution of this study is an analysis of the long-run money demand relation among OECD countries by panel unit root and cointegration testing techniques. The panel SL cointegration test and the tests of Pedroni (1999) which are also considered in the extensive simulation study of this thesis, are used to analyze the existence of stationary long-run money demand. The tests are applied on balanced panels consisting of different combinations of OECD countries comprising the period from the first quarter of 1988 to the fourth quarter of 1997. Finally, the money demand relation is estimated with the panel dynamic ordinary least squares (DOLS) method of Mark and Sul (2003).

## 1.2 Outline

The structure of the thesis is as follows: Chapter 2 provides a review of the panel cointegration tests suggested in the literature. After a brief explanation of the difference between the residual-based and maximum-likelihood-based panel cointegration tests, the derivation of the test statistics as well as their finite sample properties are summarized. In Chapter 3, by means of a Monte Carlo study, the finite sample properties of four residual-based tests of Pedroni (1999) are compared with the standardized LR-bar test of Larsson et al. (2001). Additionally, the tests are implemented to test for the existence of a cointegrating relation between the interest rate and the inflation rate of OECD countries. Chapter 4 exhibits the corrected version of the proof on the existence and the finiteness of the moments of the asymptotic trace statistic presented by Larsson et al. (2001). Furthermore, the incorrectness of the original proof provided in Larsson et al. (2001) is demonstrated. In Chapter 5, a new maximum-likelihood-based panel cointegration test is suggested. This test is an extension of the GLS-based LR trace test of Saikkonen and Lütkepohl (2000a). Additionally, the finite sample properties of this new test (panel SL test) are also summarized. In Chapter 6, the panel SL test suggested in Chapter 5 and the tests of Pedroni (1999) are applied to test the existence of a long-run money demand relation among OECD countries.

[3]This test is an extension of the cointegration test of Johansen (1995) with deterministic terms to panel data.

# Chapter 2

# Panel Cointegration Tests

Two types of panel cointegration tests can be found in the literature: residual-based tests and maximum-likelihood-based tests. Researchers introduced residual-based tests both for the null hypothesis of no cointegration and the null hypothesis of cointegration. The main idea is to test for the existence of a unit root in the residuals of a cointegrating regression equation. A unit root in the residuals implies no cointegration between the components of the model. On the contrary, the absence of a unit root in the residuals shows evidence for a cointegrating relation between the dependent and independent variables of the regression equation. Since these tests are based on the assumption that there is only one single cointegrating relation between the variables, the number of cointegrating relations cannot be detected if there are more than one.

The second type of tests are called the maximum-likelihood-based tests, which are generally extensions of the multivariate cointegration tests of Johansen (1988) to panel data. The advantage of these type of tests over the residual-based tests is that it is possible to determine the number of cointegrating relations among the variables. Moreover, the results of these tests are independent of the choice of the variable used for the necessary normalization of the cointegrating vector.

The outline of this chapter is as follows: In Section 2.1 the residual-based panel cointegration tests are examined. Among these, particular attention will be paid to the null hypothesis of no cointegration tests of Pedroni (1999, 2004), Kao (1999), Westerlund (2005a, 2006c, 2007, 2008), Westerlund and Edgerton (2007b), Hanck (2007, 2006b), Gutierrez (2008), Banerjee and Carrion-i- Silvestre (2006), Gengenbach et al. (2006), and the null hypothesis of cointegration tests of McCoskey and Kao (1998), Westerlund (2005b, 2006a,b) and Westerlund and Edgerton (2007a). In Section 2.2 the maximum-likelihood-based tests of Larsson et al. (2001), Groen and Kleibergen (2003), Breitung (2005), Anderson et al. (2006), and Larsson and Lyhagen (1999) are introduced. Section 2.3 provides a review of the finite sample properties of the presented tests.

## 2.1 Residual-Based Tests

### 2.1.1 Pedroni Tests

Pedroni (1995) introduced the first residual-based panel cointegration tests. Moreover, Pedroni (1999, 2004) extended his panel cointegration testing procedure to the case of multiple regressors.

Pedroni (1999, 2004) suggested seven different residual-based panel cointegration tests for testing the null hypothesis of no cointegration. The four within-dimension-based (i.e. panel-$v$, panel-$\rho$, semi-parametric panel-$t$ and parametric panel-$t$ ) statistics are calculated by summing up the numerator and the denominator over $N$ cross-sections separately. The three between-dimension-based (i.e. group-$\rho$, semi-parametric group-$t$ and parametric group-$t$) statistics are calculated by dividing the numerator and the denominator before summing up over $N$ cross-sections.

The starting point of the panel cointegration tests of Pedroni (1999, 2004) is the computation of the residuals of the hypothesized cointegrating regression,

$$y_{it} = \delta_{0i} + \delta_{1i}t + x_{it}'\beta_i + e_{it}, \quad i = 1, \dots, N; \quad t = 1, \dots, T, \tag{2.1}$$

in which $T$ is the number of observations over time and $N$ denotes the number of individuals in the panel. $y_{it}$ and the $K$-dimensional vector of independent variables $x_{it} = x_{i,t-1} + \nu_{it}$ are assumed to be at most $I(1)$. The cointegrating vector $\beta_i = (\beta_{1i}, \dots, \beta_{Ki})'$, the individual specific intercept $\delta_{0i}$, and the trend parameter $\delta_{1i}$ can vary over cross-sections. In addition to this, it is assumed that the error process $\mathrm{w}_{it} = (e_{it}, \nu_{it}')'$ is cross-sectionally independently identically distributed[1], and the invariance principle[2] holds individually for each

[1]The vector $\mathrm{w}_{it}$ satisfies the linear process conditions of Phillips and Solo (1992):

i. $E(\mathrm{w}_{it}\mathrm{w}_{js}') = 0$ for all $s, t, i \neq j$.

ii. $\mathrm{w}_{it} = C_i(L)\eta_{it}$, with the lag operator $L$, $C_i(L) = \sum_{j=0}^{\infty} C_{ij}L^j$ and $\sum_{j=0}^{\infty} \left\|C_{ij}C_{ij}'\right\| < \infty$. The process $\eta_{it}$ is a mean zero white noise sequence and $i.i.d.$ over $i$ and $t$ dimensions.

iii. $\Omega_i = C_i(1)C_i(1)'$.

[2]In other words, $\mathrm{w}_{it}$ fulfills, $\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]} \mathrm{w}_{it} \xrightarrow{w} B_i(\Omega_i)$ for each $i$ as $T \to \infty$ and

cross-section $i$ as $T$ grows large. Moreover, the components of $x_{it}$ should not be cointegrated among themselves[3].

The cointegrating regression in (2.1) is estimated by the ordinary least squares (OLS) method, separately for each cross-section. Additionally, the panel-$v$ and the parametric panel-$t$ statistics are calculated using the following first-differenced regression equation which is obtained by ignoring the deterministic terms.

$$\Delta y_{it} = b_{1i}\Delta x_{1it} + b_{2i}\Delta x_{2it} + \ldots + b_{Ki}\Delta x_{Kit} + \xi_{it} \tag{2.2}$$

Using the residuals from the differenced regression (2.2), the long-run variance of $\widehat{\xi}_{it}$, which is denoted by $\widehat{L}_{11i}^2$, is calculated with a Newey and West (1987) estimator.

$$\widehat{L}_{11i}^2 = \frac{1}{T}\sum_{t=1}^{T}\widehat{\xi}_{it}^2 + \frac{2}{T}\sum_{s=1}^{M_i}\left(1 - \frac{s}{M_i+1}\right)\sum_{t=s+1}^{T}\widehat{\xi}_{it}\widehat{\xi}_{i,t-s} \tag{2.3}$$

To calculate the semi-parametric statistics, the regression $\widehat{e}_{it} = \rho_i\widehat{e}_{i,t-1} + u_{it}$ is estimated using the residuals $\widehat{e}_{it}$ from the cointegrating regression (2.1). Afterwards the contemporaneous variance $(\widehat{s}_i^2)$ and the long-run variance $(\widehat{\sigma}_i^2)$ of $u_{it}$ are computed. To derive $\widehat{\sigma}_i^2$, $4\left(\frac{T}{100}\right)^{2/9}$ is used as the lag truncation function for the Newey and West kernel estimator[4]. The nearest integer is taken as the lag length for different $T$ dimensions.

The parametric test statistics, the panel-$t$ and the group-$t$, are estimated with the help of the residuals $\widehat{e}_{it}$ from the cointegrating regression (2.1), using $\widehat{e}_{it} = \rho_i\widehat{e}_{i,t-1} + \gamma_{i1}\Delta\widehat{e}_{i,t-1} + \ldots + \gamma_{i,p_i}\Delta\widehat{e}_{i,t-p_i} + u_{it}^*$. Finally, the simple variance and the long-run variance of $u_{it}^*$ are computed, which are denoted as $\widehat{s}_i^{*2}$ and $\widetilde{s}_{NT}^{*2}$, respectively. The lag truncation order of the ADF $t$-statistics can be determined using any lag order selection criterion. By the following expressions the relevant test statistics can be constructed.

a. Panel $v$-statistic

$$T^2N^{3/2}Z_{\widehat{\nu}_{NT}} \quad = \quad T^2N^{3/2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{L}_{11i}^{-2}\widehat{e}_{i,t-1}^2\right)^{-1},$$

---

$0 \leq s \leq 1$, in which $B_i(\Omega_i)$ is a Brownian motion with asymptotic covariance matrix $\Omega_i = \begin{bmatrix} \omega_{11i}^2 & \omega_{12i} \\ \omega_{21i} & \Omega_{22i} \end{bmatrix}$. The invariance principle helps to apply the functional central limit theorem.

[3]i.e. the $(K \times K)$ lower diagonal block of $\Omega_i$ is positive definite, $\Omega_{22i} > 0$.

[4]Pedroni (1995, 2004) and Newey and West (1994) recommended to use this lag truncation function.

b. Panel-$\rho$ statistic

$$T\sqrt{N}Z_{\widehat{\rho}_{NT}-1} \quad = \quad T\sqrt{N}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{e}_{i,t-1}^{2}\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{e}_{i,t-1}\Delta\widehat{e}_{it}-\widehat{\lambda}_{i}\right),$$

c. Panel-$t$ statistic (semi-parametric)

$$Z_{t_{NT}} \quad = \quad \left(\widetilde{\sigma}_{NT}^{2}\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{e}_{i,t-1}^{2}\right)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{e}_{i,t-1}\Delta\widehat{e}_{it}-\widehat{\lambda}_{i}\right),$$

d. Panel-$t$ statistic (parametric)

$$Z_{t_{NT}}^{*} = \left(\widetilde{s}_{NT}^{*2}\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{L}_{11i}^{-2}\widehat{e}_{i,t-1}^{2}\right)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{L}_{11i}^{-2}\widehat{e}_{i,t-1}\Delta\widehat{e}_{it},$$

e. Group-$\rho$ statistic

$$TN^{-1/2}\widetilde{Z}_{\widehat{\rho}_{NT}-1} \quad = \quad TN^{-1/2}\sum_{i=1}^{N}\left[\left(\sum_{t=1}^{T}\widehat{e}_{i,t-1}^{2}\right)^{-1}\sum_{t=1}^{T}\left(\widehat{e}_{i,t-1}\Delta\widehat{e}_{it}-\widehat{\lambda}_{i}\right)\right],$$

f. Group-$t$ statistic (semi-parametric)

$$N^{-1/2}\widetilde{Z}_{t_{NT}} \quad = \quad N^{-1/2}\sum_{i=1}^{N}\left[\left(\widehat{\sigma}_{i}^{2}\sum_{t=1}^{T}\widehat{e}_{i,t-1}^{2}\right)^{-1/2}\sum_{t=1}^{T}\left(\widehat{e}_{i,t-1}\Delta\widehat{e}_{it}-\widehat{\lambda}_{i}\right)\right],$$

g. Group-$t$ statistic (parametric)

$$N^{-1/2}\widetilde{Z}_{t_{NT}}^{*} = N^{-1/2}\sum_{i=1}^{N}\left[\left(\widehat{s}_{i}^{*2}\sum_{t=1}^{T}\widehat{e}_{i,t-1}^{2}\right)^{-1/2}\sum_{t=1}^{T}\widehat{e}_{i,t-1}\Delta\widehat{e}_{it}\right],$$

with

$$\widehat{\lambda}_{i} = \frac{1}{T}\sum_{s=1}^{M_i}\left(1-\frac{s}{M_i+1}\right)\sum_{t=s+1}^{T}\widehat{u}_{it}\widehat{u}_{i,t-s}, \quad \widehat{s}_{i}^{2} = \frac{1}{T}\sum_{t=1}^{T}\widehat{u}_{it}^{2},$$
$$\widehat{\sigma}_{i}^{2} = \widehat{s}_{i}^{2} + 2\widehat{\lambda}_{i}, \quad \widetilde{\sigma}_{NT}^{*2} = \frac{1}{N}\sum_{i=1}^{N}\widehat{L}_{11i}^{-2}\widehat{\sigma}_{i}^{2},$$
$$\widetilde{s}_{i}^{*2} = \frac{1}{T}\sum_{t=1}^{T}\widehat{u}_{it}^{*2}, \quad \widetilde{s}_{NT}^{*2} = \frac{1}{N}\sum_{i=1}^{N}\widetilde{s}_{i}^{*2}.$$

Note that the within-dimension-based panel-$v$ statistic is an extension of the variance ratio statistic proposed by Phillips and Ouliaris (1990). The panel-$\rho$ statistic is an extension of the semi-parametric $\rho$-statistics of Phillips and Perron (1988) and Phillips and Ouliaris (1990) to panel data. Moreover, the semi-parametric panel-$t$ statistic is a modification of the $t$-test statistic of Phillips and Perron (1988), and the parametric panel-$t$ statistic is an extension of the ADF $t$-statistic. Between-dimension-based statistics are just the group mean approach extensions of the within-dimension-based ones.

The appropriate mean and variance adjustment terms are applied, so that the test statistics are asymptotically standard normally distributed.

$$\frac{\varkappa_{NT} - m_1\sqrt{N}}{\sqrt{m_2}} \quad \stackrel{w}{\to} \quad N(0,1), \tag{2.4}$$

in which $\varkappa_{N,T}$ is the scaled form of the test statistic with respect to $N$ and $T$, and $m_1$ and $m_2$ are the moments of the underlying Brownian motion functionals. The appropriate mean and variance adjustment terms for different number of regressors and different panel cointegration test statistics are given in Table 2 in Pedroni (1999).

The null hypothesis of no cointegration for the panel cointegration test is the same for each statistic,

$$H_0 : \rho_i = 1, \text{ for all } i = 1, \ldots, N, \tag{2.5}$$

whereas the alternative hypotheses for the between-dimension-based and within-dimension-based panel cointegration tests differ. The alternative hypothesis for the between-dimension-based statistics is

$$H_1 : \rho_i < 1, \text{ for all } i = 1, \ldots, N. \tag{2.6}$$

For within-dimension-based statistics the alternative hypothesis

$$H_1 : \rho_i = \rho < 1, \text{ for all } i = 1, \ldots, N, \tag{2.7}$$

assumes a common value.

Under the alternative hypothesis, the panel-$v$ statistic diverges to positive infinity, and the right tail of the standard normal distribution is used to reject the null hypothesis. All the other panel cointegration test statistics diverge to negative infinity. Thus, the left tail of the standard normal distribution is used to reject the null hypothesis.

In his Monte Carlo study, with a data generating process (DGP) which does not allow endogeneity among the error terms, Pedroni (2004) showed that for small panels the tests are size distorted. The empirical size approaches the nominal 5% level if $T$ and/or $N$ increase. Additionally, he demonstrated that the sizes of the panel-$\rho$ and the group-$\rho$ tests converge faster to the nominal level if $N = T^{3/4}$ and $N = T^{5/6}$, respectively. To evaluate the power properties, Pedroni (1999) used an AR(1) process for the equilibrium error terms. The unadjusted powers of all the tests tend to unity as $T$ increases. When the autoregressive parameter is near unity, all tests require a larger time dimension, so that the power approaches one. Among all the tests the panel-$v$ statistic has the highest and the group-$\rho$ statistic has the lowest power.

### 2.1.2 Kao Tests

#### McCoskey and Kao (1998)

In their paper McCoskey and Kao (1998) suggested a residual-based panel cointegration test for the null hypothesis of cointegration, which is an extension of the univariate LM cointegration tests presented in Harris and Inder (1994) and Shin (1994). Let $y_{it}$ and the $K$-dimensional $x_{it}$ be integrated of order one and consider the following model.

$$\begin{aligned} y_{it} &= \alpha_i + x_{it}'\beta_i + e_{it}, && (2.8)\\ x_{it} &= x_{i,t-1} + \nu_{it}, && (2.9)\\ e_{it} &= r_{it} + u_{it}, && (2.10)\\ r_{it} &= r_{i,t-1} + \phi u_{it}, && (2.11) \end{aligned}$$

with $i = 1, \ldots, N$; $t = 1, \ldots, T$ and $u_{it} \sim (0, \sigma_u^2)$ *i.i.d.* By backward substitution the unobserved regression errors $e_{it}$ can be rewritten as a sum of a white noise and a unit root component:

$$y_{it} = \alpha_i + x_{it}'\beta_i + \phi \sum_{j=1}^{t} u_{ij} + u_{it}, \tag{2.12}$$

such that the null hypothesis of cointegration and the alternative hypothesis are formulated as

$$H_0 : \phi = 0 \quad \text{vs.} \quad H_1 : |\phi| \neq 0. \tag{2.13}$$

If $\phi = 0$, the equilibrium error process is stationary and $y_{it}$ is cointegrated with $x_{it}$. McCoskey and Kao (1998) interpreted $\phi$ as the size of the effects of a random shock on the random walk and on the stationary component.

Moreover, they assumed that $u_{it}$ and $\nu_{it}$ are weakly dependent and heterogeneously distributed innovations. Note that the error process $\mathrm{w}_{it} = (u_{it}, \nu'_{it})'$ fulfills the invariance principle and the linear process conditions of Phillips and Solo (1992) hold. Since the OLS estimator is asymptotically biased, whenever the error terms are serially correlated and $x_{it}$ is endogenous, McCoskey and Kao (1998) proposed to use alternative estimation methods: the DOLS estimator of Saikkonen (1991) and the fully-modified OLS (FMOLS) estimator of Phillips and Hansen (1990).

Thus, McCoskey and Kao (1998) proposed a panel LM test statistic based on the FMOLS estimator. They defined the long-run covariance matrix of $\mathrm{w}_{it}$ as

$$\Omega = \begin{bmatrix} \omega_1^2 & \omega_{12} \\ \omega_{21} & \Omega_{22} \end{bmatrix}. \tag{2.14}$$

The $(K \times K)$ matrix $\Omega_{22}$ is assumed to be positive definite, which means that there are no cointegrating relations among the regressors. Moreover, McCoskey and Kao (1998) introduced

$$u_{it}^{+} = u_{it} - \widehat{\omega}_{12}\widehat{\Omega}_{22}^{-1}\nu_{it}, \tag{2.15}$$

$$y_{it}^{+} = y_{it} - \widehat{\omega}_{12}\widehat{\Omega}_{22}^{-1}\nu_{it}. \tag{2.16}$$

Note that $\widehat{\Omega}_{22}$ and $\widehat{\omega}_{12}$ are consistent estimates of $\Omega_{22}$ and $\omega_{12}$, respectively. The panel LM statistic is then given as

$$\mathrm{LM}^{+} = \frac{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{j=1}^{t}\widehat{e}_{ij}^{+}\right)^2\right]}{\widehat{\omega}_{1.2}^2}, \tag{2.17}$$

in which $\widehat{e}_{it}^{+} = y_{it}^{+} - \widehat{\alpha}_i - x'_{it}\widehat{\beta}_i^{+}$ are the FMOLS residuals and $\widehat{\omega}_{1.2}^2$ is a consistent estimator of $\omega_{1.2}^2 = \omega_1^2 - \omega_{12}\Omega_{22}^{-1}\omega_{21}$ over all $T$ and $N$. $\omega_{1.2}^2$ can be estimated with a kernel estimator separately for each cross-section because the error terms are cross-sectionally independent. In (2.17) $\mathrm{LM}^{+}$ is just the average of the locally best unbiased invariant (LBUI) test statistics of Harris and Inder (1994) and Shin (1994) over cross-sections.

The panel $\mathrm{LM}^{+}$ statistic can be also computed with the DOLS estimator. Then, the test statistic $\mathrm{LM}^{+}$ should be based on the following dynamic panel regression equation.

$$y_{it} = \alpha_i + x'_{it}\beta_i + \sum_{j=-q}^{q} c_{ij}\Delta x_{i,t+j} + e_{it}^{*} \tag{2.18}$$

Thus, in (2.17) instead of $\widehat{e}_{it}^{+}$, the DOLS residuals $\widehat{e}_{it}^{*}$ from equation (2.18) are used to compute $\mathrm{LM}^{+}$.

Finally, the standardized panel LM statistic can be calculated both with FMOLS and DOLS residuals by

$$Z^{+} = \frac{\sqrt{N}(\mathrm{LM}^{+} - \mu)}{\sqrt{\sigma^{2}}}, \tag{2.19}$$

which is asymptotically $N(0,1)$ distributed as $T$ and $N \to \infty$ sequentially, i.e. $T \to \infty$ followed by $N \to \infty$. Note that $\mu$ and $\sigma^2$ are the mean and variance of the asymptotic $\mathrm{LM}^{+}$ statistic, respectively.

For the Monte Carlo study McCoskey and Kao (1998) modified the DGP designs of Harris and Inder (1994) and Phillips and Loretan (1991) to panel data. Without any serial correlation and strictly exogenous regressors, the size-adjusted power results demonstrate that the power of the panel LM test based on the FMOLS method (panel LM-FMOLS test) is higher than the power of the panel LM test based on the DOLS method (panel LM-DOLS test). This is due to the fact that the panel LM-DOLS test is generally undersized in finite samples. The powers of both tests approach unity when $T$ and $N$ increase, however $T$ should be larger than $N$. For small $T$ and $N$ both tests do not have much power. Analogous to the tests of Pedroni (2004) if there is serial correlation[5] and/or weakly exogenous error terms, both tests are size distorted. Hence, the size distortion of the panel LM-DOLS test is for most of the cases less than the size distortion of the panel LM-FMOLS test, which even approaches unity for the latter test. The size-adjusted power of the panel LM-FMOLS test is the highest when the moving average parameter is negative. On the contrary, the power of the panel LM-DOLS test is the highest if the moving average parameter is positive (or zero), and the error terms are weakly exogenous.

### Kao (1999)

Kao (1999) introduced parametric residual-based panel tests for the null hypothesis of no cointegration. He expanded the DF and ADF unit root tests to panel cointegration. The tests are based on the spurious least squares dummy variable (LSDV) panel regression equation with one single regressor.

$$y_{it} = \alpha_i + x_{it}\beta + e_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \tag{2.20}$$

in which $e_{it}$ is $I(1)$, the slope coefficient $\beta$ is assumed to be cross-section invariant (i.e. the cointegrating vector is homogeneous) and the intercept is heterogeneous. $y_{it} = \sum_{s=1}^{t} u_{is}$ and $x_{it} = \sum_{s=1}^{t} \varepsilon_{is}$ are restricted to be at

[5] For the Monte Carlo study McCoskey and Kao (1998) modelled the serial correlation by an MA(1) process.

most $I(1)$ with $u_{it} \sim (0, \sigma_u^2)$ *i.i.d.* and $\varepsilon_{it} \sim (0, \sigma_\varepsilon^2)$ *i.i.d.*. The error process $\mathrm{w}_{it} = (u_{it}, \varepsilon_{it})'$ is assumed to be independent across $i$, and it fulfills the invariance principle.

Kao (1999) proposed a DF type test using the AR(1) representation of the LSDV residuals from (2.20),

$$\widehat{e}_{it} = \rho \widehat{e}_{i,t-1} + \nu_{it}, \tag{2.21}$$

in which the AR(1) parameter $\rho$ is homogeneous. Then, the null hypothesis of no cointegration can be formulated as: $H_0 : \rho = 1$. This implies that the process $e_{it}$ is $I(1)$, i.e. $e_{it}$ has a unit root.

Based on the OLS residuals from (2.21) the estimator of $\rho$ is

$$\widehat{\rho} \;=\; \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} \widehat{e}_{it} \widehat{e}_{i,t-1}}{\sum_{i=1}^{N} \sum_{t=2}^{T} \widehat{e}_{i,t-1}^2}. \tag{2.22}$$

Kao (1999) defined

$$y_{it}^* \;=\; y_{it} - \sigma_{0u\varepsilon} \sigma_{0\varepsilon}^{-2} x_{it}, \tag{2.23}$$

$$x_{it}^* \;=\; \sigma_{0\varepsilon}^{-1} x_{it}, \tag{2.24}$$

in which $\sigma_{0\varepsilon}$ is the long-run variance of $\varepsilon_{it}$ conditional on $u_{it}$ and $\sigma_{0u\varepsilon}$ is the long-run covariance of $\varepsilon_{it}$ and $u_{it}$. Let

$$s_e^2 \;=\; \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} (\widehat{e}_{it}^* - \widehat{\rho} \widehat{e}_{i,t-1}^*)^2}{NT}, \tag{2.25}$$

with the residuals $\widehat{e}_{it}^* = y_{it}^* - \widehat{\alpha}_i^* - x_{it}^* \widehat{\beta}^*$, $\widehat{\alpha}_i^* = \widehat{\alpha}_i$ and $\widehat{\beta}^* = \sigma_{0\varepsilon} \widehat{\beta} - \sigma_{0\varepsilon}^{-1} \sigma_{0u\varepsilon}$. Note that $\widehat{\alpha}_i$ and $\widehat{\beta}$ are the LSDV estimators of $\alpha_i$ and $\beta$ from (2.20).

The DF type panel statistics that Kao (1999) proposed are

$$DF_\rho^* \;=\; \frac{\sqrt{N}T(\widehat{\rho} - 1) + \frac{3\sqrt{N}\widehat{\sigma}_\nu^2}{\widehat{\sigma}_{0\nu}^2}}{\sqrt{3 + \frac{36\widehat{\sigma}_\nu^4}{5\widehat{\sigma}_{0\nu}^4}}}, \tag{2.26}$$

$$DF_t^* \;=\; \frac{\frac{(\widehat{\rho}-1)\sqrt{\sum_{i=1}^{N} \sum_{t=2}^{T} \widehat{e}_{i,t-1}^{*2}}}{s_e} + \frac{\sqrt{6N}\widehat{\sigma}_\nu^2}{2\widehat{\sigma}_{0\nu}^2}}{\sqrt{\frac{\widehat{\sigma}_{0\nu}^2}{2\widehat{\sigma}_\nu^2} + \frac{3\widehat{\sigma}_\nu^2}{10\widehat{\sigma}_{0\nu}^2}}}. \tag{2.27}$$

Note that $\widehat{\sigma}_{0\nu}^2$ is a consistent estimator of the long-run conditional variance $\sigma_{0\nu}^2 = \sigma_{0u}^2 - \sigma_{0u\varepsilon}^2 \sigma_{0\varepsilon}^{-2}$, and $\widehat{\sigma}_\nu^2$ is a consistent estimator of the contemporaneous variance $\sigma_\nu^2 = \sigma_u^2 - \sigma_{u\varepsilon}^2 \sigma_\varepsilon^{-2}$. The term $\sigma_{0u}^2$ specifies the long-run variance of $u_{it}$, whereas $\sigma_{u\varepsilon}$ is the contemporaneous covariance between $u_{it}$ and $\varepsilon_{it}$. Contemporaneous variances can be estimated using

$$\widehat{\Omega} = \begin{pmatrix} \widehat{\sigma}_u^2 & \widehat{\sigma}_{u\varepsilon} \\ \widehat{\sigma}_{u\varepsilon} & \widehat{\sigma}_\varepsilon^2 \end{pmatrix} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{\mathrm{w}}_{it}\widehat{\mathrm{w}}_{it}'}{NT}. \tag{2.28}$$

The estimation of the long-run variance and covariances requires to choose an appropriate bandwidth and a kernel estimator. Under the null hypothesis the DF type statistics are asymptotically standard normally distributed as $T \to \infty$ followed by $N \to \infty$.

In contrast to (2.21), the ADF type panel statistic is based on the following AR($p$) regression.

$$\widehat{e}_{it} = \rho\widehat{e}_{i,t-1} + \gamma_1\Delta\widehat{e}_{i,t-1} + \ldots + \gamma_p\Delta\widehat{e}_{i,t-p} + \nu_{itp}, \tag{2.29}$$

in which $\widehat{e}_{it}$ depends also on the lagged changes of the LSDV residuals. Kao (1999) formulated the ADF panel test statistic as

$$ADF = \frac{\frac{\sum_{i=1}^{N}(e_i'Q_i\nu_i)}{s_\nu\sqrt{\left[\sum_{i=1}^{N}(e_i'Q_ie_i)\right]}} + \frac{\sqrt{6N}\widehat{\sigma}_\nu}{2\widehat{\sigma}_{0u}}}{\sqrt{\frac{\widehat{\sigma}_{0\nu}^2}{2\widehat{\sigma}_\nu^2} + \frac{3\widehat{\sigma}_\nu^2}{10\widehat{\sigma}_{0\nu}^2}}}, \tag{2.30}$$

with

$$Q_i = I - X_{ip}(X_{ip}'X_{ip})^{-1}X_{ip}', \tag{2.31}$$

and $X_{ip}$ denoting a matrix of observations on the $p$ regressors ($\Delta\widehat{e}_{i,t-1}, \Delta\widehat{e}_{i,t-2}$ $, \ldots, \Delta\widehat{e}_{i,t-p}$). Note that $e_i$ is the vector of observations on $\widehat{e}_{i,t-1}$, and

$$s_\nu^2 = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{\nu}_{itp}^2}{NT}, \tag{2.32}$$

with $\widehat{\nu}_{itp}$ being the estimate of $\nu_{itp}$. Under the null hypothesis the panel ADF test of Kao (1999) is also asymptotically $N(0,1)$ distributed as $T$ and $N \to \infty$ sequentially.

If the $x_{it}$ regressors are not cointegrated, then the tests can be also implemented to multiple regressors case.

To find out the finite sample properties of the tests, Kao (1999) employed a Monte Carlo study. He added the following two bias corrected test statistics to the simulation study.

$$DF_\rho = \frac{\sqrt{N}T(\widehat{\rho}-1) + 3\sqrt{N}}{\sqrt{51/5}}, \tag{2.33}$$

$$DF_t = \sqrt{\frac{5t_\rho}{4}} + \sqrt{\frac{15N}{8}}, \tag{2.34}$$

in which $t_\rho = \frac{(\widehat{\rho}-1)\sqrt{\sum_{i=1}^{N}\sum_{t=2}^{T}\widehat{e}_{i,t-1}^{*2}}}{s_e}$ is the $t$-statistic for $\rho = 1$. He constituted the simulation study on the modified version of the DGP used by Engle and Granger (1987) and Gonzalo (1994). All the five tests based on the DGP with endogenous regressors have size distortions. However, the empirical sizes of the $DF_t^*$ and $DF_\rho^*$ tests approach the 5% significance level for large $T$ and $N$. If $T$ and $N$ are small, none of the tests have high power, and $DF_t^*$ and $DF_\rho^*$ show the lowest power as expected. Hence, among all the tests $DF_\rho$ and $DF_t$ have the best power, and the powers of all the tests increase with $T$ and $N$. Moreover, the tests are size distorted and lose power in the presence of serial correlation and different levels of endogeneity in the error terms. Overall, $DF_t^*$ and $DF_\rho^*$ tests have better size and power properties[6].

### 2.1.3 Westerlund Tests

**Westerlund (2005a)**

Westerlund (2005a) proposed two residual-based panel cointegration tests for the null hypothesis of no cointegration. He extended the univariate variance ratio tests of Breitung (2002) to panel data. Both of these statistics are non-parametric, and they require neither a specification for the underlying DGP nor the estimation of the nuisance parameters.

The non-parametric tests have significant advantages over the parametric and semi-parametric relatives. First, they are easy to compute as the number of calculations needed are relatively less compared to those used in the parametric and semi-parametric tests. Second, it is not necessary to correct for the effects of the dependent data, which prevents nuisance parameter problem. The parametric tests face instead the problem of selecting the right lag order of the autoregressive process, and the semi-parametric tests require the right bandwidth selection for the kernel estimator. Since the lag and bandwidth selection should be done for each cross-section separately, the problem is much more serious within panel data framework. As a result, the parametric and semi-parametric test statistics depend on the true DGP and the right truncation parameter. For his panel cointegration tests Westerlund (2005a) considered the heterogeneous panel regression equation (2.1) with the assumption that $y_{it}$ and the $K$-dimensional vector of regressors $x_{it} = x_{i,t-1} + \nu_{it}$ are at most $I(1)$. The error process $\mathrm{w}_{it} = (e_{it}, \nu_{it}')'$ is cross-

[6]Please note that the power results are not size-adjusted.

sectionally independent[7] and has an MA($\infty$) representation[8]. Moreover, the invariance principle holds for $\mathrm{w}_{it}$ individually for each cross-section $i$ as $T$ grows. It is also assumed that the components of $x_{it}$ are not cointegrated among themselves.

Westerlund (2005a) proposed the following variance ratio test statistics:

a. Panel variance ratio statistic

$$VR_P = \sum_{i=1}^{N}\sum_{t=1}^{T}\left(\sum_{j=1}^{t}\widehat{e}_{ij}\right)^2\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{e}_{it}^2\right)^{-1}. \tag{2.35}$$

b. Group variance ratio statistic

$$VR_G = \sum_{i=1}^{N}\left[\sum_{t=1}^{T}\left(\sum_{j=1}^{t}\widehat{e}_{ij}\right)^2\left(\sum_{t=1}^{T}\widehat{e}_{it}^2\right)^{-1}\right]. \tag{2.36}$$

The null hypothesis is based on the following autoregressive process.

$$\widehat{e}_{it} = \rho_i\widehat{e}_{i,t-1} + u_{it}, \tag{2.37}$$

in which $\widehat{e}_{it}$'s are the OLS residuals from (2.1) estimated separately for each cross-section. The null hypothesis of no cointegration is tested by checking whether a unit root is present in the residuals of regression (2.1). This can be represented as

$$H_0 : \rho_i = 1, \quad \text{for all } i. \tag{2.38}$$

For the two different panel cointegration statistics, Westerlund (2005a) formulated two different alternative hypotheses. The alternative hypothesis for the panel variance ratio statistic is formulated as

$$H_1 : |\rho_i| = |\rho| < 1, \quad \text{for all } i. \tag{2.39}$$

On the contrary, the alternative hypothesis for the group variance ratio statistic does not assume that all the cross-sections in the panel data are cointegrated, but allows for a fraction of the panel data to be cointegrated. This can be denoted as

$$H_1 : |\rho_i| < 1, \ \text{for } i = 1, \ldots, N_1 \text{ and } \rho_i = 1, \ \text{for } i = N_1 + 1, \ldots, N, \tag{2.40}$$

---

[7]If the assumption about the independence of the individuals is violated, Westerlund (2005a) proposed to use the demeaned version of $x_{it}$ and $y_{it}$ with respect to the common time effects.

[8]Please refer to Section 2.1.1 for further assumptions about $\mathrm{w}_{it}$.

with $N_1/N \rightarrow \psi$, $\psi \in (0,1]$ as $N \rightarrow \infty$.

With the sequential limit theorem of Phillips and Moon (1999), Westerlund (2005a) proved that under the null hypothesis the standardized panel and group variance ratio statistics are asymptotically standard normally distributed as $T$ and $N$ go to infinity sequentially. Therefore, the first two moments of the asymptotic panel and group variance ratio test statistics are used for the standardization. Since the asymptotic panel and group variance ratio test statistics depend on the Brownian motion functionals, Westerlund (2005a) simulated their first two moments for different number of regressors and deterministic specifications.

To reject the null hypothesis, the left tail of the standard normal distribution is used because under the alternative hypothesis the standardized panel statistics converge to negative infinity. Note that the finite sample properties of the tests are explained in Section 2.3.

**Westerlund (2005b)**

Based on the regression equation (2.1) Westerlund (2005b) suggested a residual-based panel cointegration test for the null hypothesis of cointegration, which accommodates for mixtures of cointegrated and spurious regressions. Moreover, the assumptions for $y_{it}$, $x_{it}$ and the error process $\mathrm{w}_{it}$ are same as in Westerlund (2005a). Westerlund (2005b) extended the time series CUSUM tests of Xiao (1999) and Xiao and Phillips (2002) to panel data. He considered three different models: Model 1: $\delta_{0i} = \delta_{1i} = 0$, Model 2: $\delta_{0i} \neq 0$, $\delta_{1i} = 0$, and Model 3: $\delta_{0i} \neq 0$, $\delta_{1i} \neq 0$.

If $y_{it}$ and $K$-dimensional vector of regressors $x_{it}$ in (2.1) are cointegrated, then the residuals should be $I(0)$. Thus, the fluctuations should be long-run equilibrium errors.

To overcome the nuisance parameters problem caused by the endogeneity of the regressors and the serial correlation of the error terms, the residuals of (2.1) are estimated either with the FMOLS estimator of Phillips and Hansen (1990) or the DOLS estimators of Saikkonen (1991) and Stock and Watson (1993). Because in the presence of endogeneity and serial correlation, both of the estimators are unbiased and asymptotically efficient, whereas the OLS estimator is asymptotically biased and inefficient.

The null hypothesis of the panel CUSUM test is that the whole panel is cointegrated, and under the alternative hypothesis a fraction of the panel is not cointegrated. Suppose $N_1$ is the number of individual DOLS or FMOLS residuals ($\widehat{e}^*_{it}$), which have a unit root, i.e. $y_{it}$ and $x_{it}$ are not cointegrated. Using the relation $N_1/N \rightarrow \psi$, $\psi \in (0,1]$ as $N \rightarrow \infty$, the null and alternative hypotheses can be formulated as

$$H_0 : \psi = 0 \quad \text{vs.} \quad H_1 : \psi > 0. \tag{2.41}$$

The null hypothesis of cointegration can be tested by looking at the fluctuations of $\widehat{e}^*_{it}$. The panel CUSUM test statistic can be computed by

$$CS_{NT} = \frac{1}{N} \sum_{i=1}^{N} \left( \max_{t=1,\ldots,T} \frac{1}{\widehat{\omega}_{i1.2} T^{1/2}} |S^*_{it}| \right), \tag{2.42}$$

with $S^*_{it} = \sum_{j=1}^{t} \widehat{e}^*_{ij}$ and $\widehat{\omega}_{i1.2}$ being any consistent semi-parametric kernel estimator of the long-run variance of $e_{it}$ conditional on $\nu_{it}$. $CS_{NT}$ is just the average of the individual test statistic of Xiao and Phillips (2002). The choice of the appropriate kernel estimator and the bandwidth parameter is crucial for finding a consistent estimate of $\omega_{i1.2}$. Westerlund (2005b) chose the Bartlett kernel as the kernel function because it ensures the non-negativity of the long-run covariance estimates. He selected a fixed bandwidth parameter, i.e. $[T^{(1/3)}]$, so that the test has asymptotic power against the null hypothesis.

With the sequential limit theorem Westerlund (2005b) demonstrated that under the null hypothesis the standardized CUSUM test statistic[9] ($Z_{NT}$) is asymptotically standard normally distributed as $T \to \infty$ followed by $N \to \infty$, whereas under the alternative hypothesis the statistic converges to positive infinity. In other words, the decision of the test is made on the right tail of the standard normal distribution.

### Westerlund (2006a)

In this study Westerlund (2006a) proposed a test procedure, which reduces the size distortions of the panel LM cointegration test of McCoskey and Kao (1998). As I pointed out earlier in Section 2.1.2, McCoskey and Kao (1998) proposed the panel LM test for the null hypothesis of cointegration. In his Monte Carlo study, Westerlund (2005b) demonstrated that the test of McCoskey and Kao (1998) has severe size distortions, whenever the autoregressive parameter of the error process is close to unity. However, Westerlund (2005b) showed that in general the panel LM test of McCoskey and Kao (1998) has higher power than his panel CUSUM test. This led Westerlund (2006a) to propose a new testing procedure to eliminate the size distortions of the test of McCoskey and Kao (1998).

[9]The simulated first two moments for the standardization of the panel CUSUM test statistic can be found in Westerlund (2005b), Table I. The moments are tabulated for three different models and $K = 1, \ldots 4$.

The straightforward procedure that Westerlund (2006a) presented to diminish the size distortions is established on the study of Choi (2004). The idea is to break the sample into two subsamples, in which one subsample consists of the odd numbered observations and the other one of the even numbered ones. The standardized panel LM test statistics are computed for each subsample separately, which are named as $Z_1^+$ and $Z_2^+$ for the first and the second subsample, respectively. Finally, the two test statistics are combined to $Z_M^+ \equiv \max\{Z_1^+, Z_2^+\}$ using the Bonferroni principle. The Bonferroni inequality

$$P(Z_M^+ > Z_{\alpha/2}) \leq \alpha, \tag{2.43}$$

with $Z_{\alpha/2}$ being the $\alpha/2$ level critical value from the standard normal distribution, states that the significance level of $Z_M^+$ is constrained from above by $\alpha$. In a Monte Carlo study Westerlund (2006a) compared his test with the panel cointegration test of McCoskey and Kao (1998). The Monte Carlo study is based on FMOLS and DOLS residuals from a single regressor DGP with a constant intercept. The $Z_M^+$ test has lower size distortions than the standardized panel LM test when the autoregressive parameter of the error term is high. Moreover, there is no sign of loss in power. Both tests are less size distorted if the statistics are derived with FMOLS method. Note that the test procedure of Westerlund (2006a) can also be applied to other panel cointegration tests.

### Westerlund (2006b)

As pointed out in Chapter 1, the panel cointegration tests which do not allow for structural breaks are size distorted, in the presence of such breaks. Building on this information Westerlund (2006b) extended the test of McCoskey and Kao (1998) and proposed a test which allows for multiple structural breaks in the deterministic terms. The test can be implemented if the number of breaks is unknown and their locations may be different for each cross-section. In addition, the test allows for endogenous regressors and serial correlation of the error terms. The panel LM test with multiple structural breaks is based on the following DGP (cp. (2.8)-(2.11)).

$$\begin{aligned} y_{it} &= d_{it}'\delta_{ij} + x_{it}'\beta_i + e_{it}, && (2.44)\\ x_{it} &= x_{i,t-1} + \nu_{it}, && (2.45)\\ e_{it} &= r_{it} + u_{it}, && (2.46)\\ r_{it} &= r_{i,t-1} + \phi_i u_{it}, && (2.47) \end{aligned}$$

with $i = 1, \ldots, N$ and $t = 1, \ldots, T$. The assumptions on the error process $\mathrm{w}_{it} = (u_{it}, \nu_{it}')'$ are identical to those made in McCoskey and Kao (1998).

$d_{it}$ represents the vector of deterministic terms. $\beta_i$ and $\delta_{ij}$ are the unknown coefficient vectors. $j$ denotes the index for the structural breaks, and there are at most $M_i$ structural breaks for each cross-section at dates $T_{i1}, \ldots, T_{iM_i}$ with $T_{i0} = 1$ and $T_{iM_i+1} = T$. The initial value of $r_{it}$ is set to zero. To derive the test statistic Westerlund (2006b) assumed that the number and the dates of the breaks are known. The dates of the structural breaks are defined as $T_{ij} = [\lambda_{ij} T]$ for $j = 1, \ldots, M_i$, such that $\lambda_{ij} \in (0,1)$. He considered five models with different deterministic specifications: Model 1: $d_{it} = \{\varnothing\}$, Model 2: $d_{it} = 1$, Model 3: $d_{it} = (1,t)'$, Model 4: $d_{it} = 1$ and $M_i > 0$ for at least one $i$, and Model 5: $d_{it} = (1,t)'$ and $M_i > 0$ for at least one $i$.

Using $r_{i0} = 0$ and the stationarity of $u_{it}$, the null hypothesis of cointegration is formulated as

$$H_0 : \phi_i = 0, \text{ for all } i = 1, \ldots, N \text{ vs.} \tag{2.48}$$
$$H_1 : \phi_i \neq 0, \text{ for } i = 1, \ldots, N_1 \text{ and } \phi_i = 0, \text{ for } i = N_1 + 1, \ldots, N. \tag{2.49}$$

The null hypothesis states that all individuals in the panel are cointegrated, whereas under the alternative only some individuals of the panel is cointegrated. In contrast to McCoskey and Kao (1998), $\phi_i$ can differ across individuals. In addition to the assumptions above, suppose that $N_1/N \to \psi$ as $N \to \infty$ for $\psi \in (0,1]$. Westerlund (2006b) proposed the following panel LM statistic

$$Z(M) = \sum_{i=1}^{N} \sum_{j=1}^{M_i+1} \sum_{t=T_{i,j-1}+1}^{T_{ij}} (T_{ij} - T_{i,j-1})^{-2} \widehat{\omega}_{i1.2}^{-2} \left( \sum_{k=T_{i,j-1}+1}^{t} \widehat{e}_{ik}^{*} \right)^2 . \tag{2.50}$$

$\widehat{e}_{it}^{*}$ can be either DOLS or FMOLS residuals. To obtain an efficient estimator $\widehat{e}_{it}^{*}$ of the error terms $e_{it}$, for each subsample $j$ ($j = 1, \ldots, M_i$) of individual $i$, the model is estimated separately from $T_{i,j-1}$ to $T_{ij}$ time series observations. $\widehat{\omega}_{i1.2}^2$ is a consistent semi-parametric kernel estimator of the long-run variance of $u_{it}$ conditional on $\nu_{it}$. For this purpose, Westerlund (2006b) suggested to use the Bartlett kernel as weight function and $[T^{1/3}]$ as the bandwidth parameter. Since $\omega_{i1.2}^2$ is assumed to be constant, the full length of the time series dimension can be used for estimation.

Under the null hypothesis the test statistic has a limiting normal distribution free of nuisance parameters as $T$ and $N$ go to infinity sequentially. Since Westerlund (2006b) showed that the limiting distribution does not depend on the location and the number of structural breaks, the test is easy to implement. As structural breaks are allowed under both, the null and alternative hypotheses, rejection of the null hypothesis does not necessarily mean that there are no structural breaks.

To obtain the number and the location of the breaks, which are unknown, Westerlund (2006b) used the two-step procedure of Bai and Perron (1998, 2003). In this procedure the locations and the number of the breaks are achieved by globally minimizing the sum of squared residuals.

$$\widehat{T}_i = \arg\min_{T_i} \sum_{j=1}^{M_i+1} \sum_{t=T_{i,j-1}+1}^{T_{ij}} (y_{it} - d'_{it}\widehat{\delta}_{ij} - x'_{it}\widehat{\beta}_i)^2, \tag{2.51}$$

in which $\widehat{T}_i = (\widehat{T}_{i1}, \ldots, \widehat{T}_{iM_i})'$ is the vector of estimated break points.

In a first step, the unknown regression parameters are estimated together with the unknown break points using the full length time series observations for each cross-section. The minimization of the sum of squared residuals is performed iteratively by dynamic programming. Given $\beta_i$ as starting value, the objective function is minimized with respect to $\delta_{ij}$ and $T_i$. Next, holding $T_i$ fixed, minimization is conducted with respect to $\beta_i$ and $\delta_{ij}$. The iteration continues until the objective function reaches its global minimum.

The first step delivers the estimated break partitions and the sum of squared residuals for each number of breaks that lies in the interval $[1, J]$, with $J$ being a predetermined upper bound for the possible number of breaks.

In a second step, the number of breaks are estimated by an information criterion[10]. The two steps are repeated $N$ times and finally, LM statistic is calculated using $\widehat{T}_i$. The limiting distribution of the panel LM statistic with structural breaks does not change if the number and the locations of the breaks for each individual is unknown. From a Monte Carlo study, Westerlund (2006b) concluded that $Z(M)$ has acceptable power and small size distortions. However, the test is severely size distorted if the individual subsamples before or after the break are not long enough. Moreover, the power of the test decreases if the break dates are unknown. The break estimates are more accurate for large $T$. The accuracy of the estimates depend also on the magnitude of the breaks.

**Westerlund (2006c)**

In a more recent study, Westerlund (2006c) presented four simple tests for the null hypothesis of no cointegration that allow for a time varying cointegrating relation under both the null and alternative hypotheses. The tests are extensions of the univariate time series tests developed by Gregory and Hansen (1996a). The new panel tests allow for a single unknown break in the level, which is located at different dates for different individuals. The

[10] Westerlund (2006b) recommended the Schwarz Bayesian information criterion.

assumptions for $y_{it}$, $x_{it}$ and the error process $\mathrm{w}_{it}$ are the same as in Westerlund (2005a). The test statistic is based on the residuals of the following least squares regression with a level shift.

$$y_{it} = \delta_{0i} + \delta_{1i}t + \eta_i D_{it} + x'_{it}\beta + e_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T \quad (2.52)$$

Suppose that the location of the shift, which is denoted by $T_i$ for individual $i$, can be defined as $T_i = [\lambda_i T]$ with $\lambda_i \in (0, 1)$. Then, the shift dummy variable $D_{it}$ is 1 if there is a shift for individual $i$ for $t > T_i$, and $D_{it} = 0$, for $t \leq T_i$.

Westerlund (2006c) considered two models with different deterministic specifications: In Model 1 there is only a heterogeneous intercept ($\delta_{0i} \neq 0$, $\delta_{1i} = 0$), and in Model 2 there is a heterogeneous intercept and a linear time trend ($\delta_{0i} \neq 0$, $\delta_{1i} \neq 0$).

Using an AR(1) representation of the regression residuals, i.e. $\widehat{e}_{it} = \rho_i \widehat{e}_{i,t-1} + u_{it}$ with $u_{it} \sim (0, \sigma_u^2)$ $i.i.d$, the null and alternative hypotheses are formulated. Thus, $H_0 : \rho_i = 1$ for all $i$ is tested against $H_i : |\rho_i| < 1$ for at least one $i$. In other words, the null hypothesis of no cointegration is tested by testing the regression residuals for a unit root. If $\widehat{e}_{it}$, $i = 1, \ldots, N$, has a unit root, then $y_{it}$ and $x_{it}$ are not cointegrated.

Westerlund (2006c) defined the following semi-parametric test statistics:

$$SZ_t = \sum_{i=1}^{N} \min_{\lambda_i} \widehat{\omega}_i^{-1} \left( \sum_{t=1}^{T} \widehat{e}_{i,t-1}^2 \right)^{-1/2} \left( \sum_{t=1}^{T} \widehat{e}_{i,,t-1} \Delta \widehat{e}_{it} - T\widehat{\tau}_i \right), \quad (2.53)$$

$$SZ_\rho = \sum_{i=1}^{N} \min_{\lambda_i} \left( \sum_{t=1}^{T} \widehat{e}_{i,t-1}^2 \right)^{-1} \left( \sum_{t=1}^{T} \widehat{e}_{i,t-1} \Delta \widehat{e}_{it} - T\widehat{\tau}_i \right), \quad (2.54)$$

in which $\hat{\omega}_i^2 = \sum_{s=-M}^{M} \left(1 - \frac{s}{1-M}\right) \widehat{\gamma}_{is}$, $\widehat{\tau}_i = \sum_{s=1}^{M} \left(1 - \frac{s}{1-M}\right) \widehat{\gamma}_{is}$, and $\widehat{\gamma}_{is}$ denotes the $s$th order autocovariance of the least square estimates $\widehat{u}_{it}$ of $u_{it}$.

Using $\widehat{e}_{it}^* = \widehat{e}_{it} - \widehat{\alpha}_{i1} \Delta \widehat{e}_{i,t-1} - \ldots - \widehat{\alpha}_{ip} \Delta \widehat{e}_{i,t-p}$ and $\widehat{e}_{it}^* = \rho_i \widehat{e}_{i,t-1}^* + u_{it}^*$, Westerlund (2006c) defined the parametric test statistics

$$PZ_t = \sum_{i=1}^{N} \min_{\lambda_i} \widehat{\gamma}_{i0}^{*-1} \left( \sum_{t=1}^{T} \widehat{e}_{i,t-1}^{*2} \right)^{-1/2} \left( \sum_{t=1}^{T} \widehat{e}_{i,t-1}^* \Delta \widehat{e}_{it}^* \right), \quad (2.55)$$

$$PZ_\rho = \sum_{i=1}^{N} \min_{\lambda_i} \left( \sum_{t=1}^{T} \widehat{e}_{i,t-1}^{*2} \right)^{-1} \left( \sum_{t=1}^{T} \widehat{e}_{i,t-1}^* \Delta \widehat{e}_{it}^* \right). \quad (2.56)$$

Here $\widehat{\gamma}_{i0}^*$ is the variance of the least squares estimates $\widehat{u}_{it}^*$ of $u_{it}^*$. Applying the sequential limit theorem, Westerlund (2006c) showed that under the null hypothesis the standardized versions of these four test statistics with respect to $T$ and $N$, have a limiting normal distribution. These statistics are standardized using the appropriate asymptotic moments[11]. The test decision is made on the left tail of the standard normal distribution. According to his simulation results all the tests are size distorted, and $SZ_t$ and $PZ_\rho$ often overreject the true null hypothesis. On the contrary, for most of the cases $SZ_\rho$ is undersized. Among all the tests the size of $PZ_t$ is the closest to the 5% nominal level. If the autoregressive parameter of the equilibrium error terms is near unity, the tests have difficulties rejecting the false null hypothesis, especially for small $T$ and $N$.

**Westerlund (2007)**

Westerlund (2007) proposed four residual-based panel cointegration tests for the null hypothesis of no cointegration, which are generalized versions of the tests proposed in Banerjee et al. (1998). In contrast to e.g. Pedroni (1999, 2004), no common factor restrictions[12] are imposed because if the restriction is violated the tests lose power.

The tests of Westerlund (2007) accommodate serially correlated error terms and non-strictly exogenous regressors. The tests are based on the following error correction model.

$$\begin{aligned} y_{it} &= \delta_{0i} + \delta_{1i} t + z_{it}, && (2.57) \\ x_{it} &= x_{i,t-1} + \nu_{it}, && (2.58) \end{aligned}$$

with $i = 1, \ldots, N$; $t = 1, \ldots, T$. The $K$-dimensional vector $x_{it}$ is $I(1)$. In addition to this, $z_{it}$ is modelled as

$$\alpha_i(L)\Delta z_{it} = \alpha_i(z_{i,t-1} - \beta_i' x_{i,t-1}) + \gamma_i(L)'\nu_{it} + e_{it}. \tag{2.59}$$

$\alpha_i$ denotes the error correction term, $\alpha_i(L) = 1 - \sum_{j=1}^{p_i} \alpha_{ij} L^j$ and $\gamma_i(L) = \sum_{j=0}^{p_i} \gamma_{ij} L^j$ with $L$ as the lag operator. In order to get the conditional error correction model for $y_{it}$, Westerlund (2007) replaced $z_{it}$ in (2.59) using (2.57), and obtained

$$\alpha_i(L)\Delta y_{it} = \phi_{0i} + \phi_{1i} t + \alpha_i(y_{i,t-1} - \beta_i' x_{i,t-1}) + \gamma_i(L)'\nu_{it} + e_{it}. \tag{2.60}$$

The deterministic terms are defined as $\phi_{0i} = \alpha_i(1)\delta_{1i} - \alpha_i\delta_{0i} + \alpha_i\delta_{1i}$ and $\phi_{1i} = -\alpha_i\delta_{1i}$. Westerlund (2007) considered three different models with different

[11] The simulated asymptotic moments can be found in Westerlund (2006c), Table 1.

[12] Westerlund (2007) defined the common factor restriction as the assumption that the long-run cointegrating vector of the variables in their levels is equal to the short-run adjustment process of the variables in their differences.

deterministic specifications. In Model 1 there are no deterministic terms, i.e. $\delta_{0i} = \delta_{1i} = 0$, in Model 2 there is only an intercept, i.e. $\delta_{0i} \neq 0, \delta_{1i} = 0$ and in Model 3 there is an intercept and a linear time trend, i.e. $\delta_{0i} \neq 0, \delta_{1i} \neq 0$. The assumptions on the error process $\mathrm{w}_{it} = (e_{it}, \nu'_{it})'$ and the $K$-dimensional vector $x_{it}$ are the same as the assumptions in Westerlund (2005a) presented above. Moreover, it is assumed that $x_{it}$ is weakly exogenous. This assumption enables us to test for no cointegration using (2.60).

Based on the conditional error correction model (2.60) the null hypothesis of no cointegration can be formulated as

$$H_0 : \alpha_i = 0, \quad \text{for all } i. \tag{2.61}$$

The alternative hypothesis for the two panel statistics [13] is

$$H_1 : \alpha_i = \alpha < 0, \quad \text{for all } i, \tag{2.62}$$

in which a common error correction parameter $\alpha$ is assumed for all cross-sections. Thus, the rejection of the null hypothesis emphasizes that the whole panel is cointegrated with the assumption that $\nu_{it}$ and $e_{it}$ are stationary. For the two group mean statistics the alternative hypothesis is

$$H_1 : \alpha_i < 0, \quad \text{for at least one } i, \tag{2.63}$$

in which there is no common value for the error correction parameter, and the rejection of the null hypothesis means that for at least one cross-section $y_{it}$ and $x_{it}$ are cointegrated.

The tests are developed using the reparameterized version of (2.60), which is

$$\Delta y_{it} = \phi'_i d_t + \alpha_i y_{i,t-1} + \lambda'_i x_{i,t-1} + \sum_{j=1}^{p_i} \alpha_{ij} \Delta y_{i,t-j} + \sum_{j=0}^{p_i} \gamma_{ij} \Delta x_{i,t-j} + e_{it}, \tag{2.64}$$

$\phi_i = (\phi_{0i}, \phi_{1i})'$, $d_t = (1, t)'$ and $\lambda_i = -\alpha_i \beta_i$.

To derive the two panel statistics, first the lag order $p_i$ for each cross-section should be determined. Afterwards, $\Delta y_{it}$ and $y_{i,t-1}$ are regressed on $d_t$, on the lags of $\Delta y_{it}$, as well as on the contemporaneous and lagged $\Delta x_{it}$. Using the residuals from these first and second regressions, which are symbolized by $\Delta \widetilde{y}_{it}$ and $\widetilde{y}_{i,t-1}$, respectively, the common error correction parameter $\alpha$ and its standard error $\sigma_\alpha$ are estimated. Hence, the estimators of $\alpha$ and $\sigma_\alpha$

[13] i.e. pooling along the cross-section dimension.

are

$$\widehat{\alpha} = \left(\sum_{i=1}^{N}\sum_{t=2}^{T}\widetilde{y}_{i,t-1}\right)^{-1}\sum_{i=1}^{N}\sum_{t=2}^{T}\frac{1}{\widehat{\alpha}_i(1)}\widetilde{y}_{i,t-1}\Delta\widetilde{y}_{it}, \tag{2.65}$$

$$\widehat{\sigma}_{\widehat{\alpha}} = \left[\left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\widehat{\sigma}_i}{\widehat{\alpha}_i(1)}\right)^2\right)^{-1}\sum_{i=1}^{N}\sum_{t=2}^{T}\widetilde{y}_{i,t-1}^2\right]^{-1/2}. \tag{2.66}$$

The term $\widehat{\sigma}_i$ is the estimated standard error obtained by applying OLS to (2.64). Thus, the panel statistics are

$$P_\tau = \frac{\widehat{\alpha}}{\widehat{\sigma}_{\widehat{\alpha}}}, \tag{2.67}$$

$$P_\alpha = T\widehat{\alpha}. \tag{2.68}$$

To compute the two group mean statistics, the parameters of (2.64) are estimated with OLS for each cross-section separately. Note that the cross-section variant lag order $p_i$ can be determined by any information criterion or by a top-down procedure. Next, $\alpha_i(1)$ can be estimated with the help of the fact that under the null hypothesis $\omega_{yi}^2 = \frac{\omega_{ui}^2}{\alpha_i^2(1)}$, where $\omega_{yi}^2$ is the long-run variance of $\Delta y_{it}$, and $\omega_{ui}^2$ is the long-run variance of the composite error term $u_{it} = \gamma_i(L)'\nu_{it} + e_{it}$. Hence, $\widehat{\alpha}_i(1) = \frac{\widehat{\omega}_{ui}^2}{\widehat{\omega}_{yi}^2}$, with $\widehat{\omega}_{ui}^2$ and $\widehat{\omega}_{yi}^2$ being the kernel estimators of the long-run variances $\omega_{ui}^2$ and $\omega_{yi}^2$, respectively. In this estimation procedure for $\alpha_i$, the bandwidth parameter selection problem occurs. If models 2 and 3 are under consideration, the estimation of $\omega_{yi}^2$ using a kernel estimator instead of $\Delta y_{it}$, requires the usage of the fitted residuals from a first-stage regression of $\Delta y_{it}$ on $d_t$.

Westerlund (2007) formulated the group mean test statistics as

$$G_\tau = \frac{1}{N}\sum_{i=1}^{N}\frac{\widehat{\alpha}_i}{\widehat{\sigma}_{\widehat{\alpha}_i}}, \tag{2.69}$$

$$G_\alpha = \frac{1}{N}\sum_{i=1}^{N}\frac{T\widehat{\alpha}_i}{\widehat{\alpha}_i(1)}, \tag{2.70}$$

in which $\widehat{\sigma}_{\widehat{\alpha}_i}$ is the standard error of $\widehat{\alpha}_i$.

Under the null hypothesis all the four test statistics have limiting normal distribution as $T$ and $N$ go to infinity sequentially. In other words the statistics are standard normally distributed when standardized with appropriate moments. Hence, the asymptotic distributions and the moments are dependent on the deterministic terms and the number of regressors included in the regression model.

Under the alternative hypothesis the group mean statistics $G_\tau$, $G_\alpha$ and the panel statistics $P_\tau$, $P_\alpha$ diverge to negative infinity, which means that the test decision is made on the left tail of the standard normal distribution.

Additionally, Westerlund (2007) proposed to use the bootstrap approach of Chang (2004) for panel cointegration testing, which accommodates for cross-sectional dependence.

**Westerlund (2008)**

In this study, Westerlund (2008) proposed two panel cointegration tests, which are powerful for testing the Fisher Effect[14]. The tests allow for cross-sectional dependence and do not require *a priori* knowledge about the integratedness of the variables. Under the assumption that the Fisher Effect holds, the tests are constructed on the following equations.

$$n_{it} = \alpha_i + \beta_i \pi_{it} + z_{it}, \tag{2.71}$$
$$\pi_{it} = \delta_i \pi_{i,t-1} + \nu_{it}. \tag{2.72}$$

$n_{it}$ represents the *ex post* nominal interest rate at time $t$ for country $i$ and $\pi_{it}$ is the actual nominal interest rate at time $t$ for country $i$. The inflation rate can be either nonstationary ($\delta_i = 1$) or stationary ($\delta_i < 1$). The error process $z_{it}$ allows for cross-sectional dependence, which is modelled as

$$z_{it} = \lambda_i' F_t + e_{it}, \tag{2.73}$$
$$F_{jt} = \rho_j F_{j,t-1} + u_{jt}, \tag{2.74}$$
$$e_{it} = \vartheta_i e_{i,t-1} + \xi_{it}. \tag{2.75}$$

$F_t$ is the $m$-dimensional vector of common factors[15] $F_{jt}$ for $j = 1, \ldots, m$ and $\lambda_i$ is the corresponding vector of factor loadings. Strict stationarity of $F_t$ is ensured if $\rho_j < 1$ for all $j$. Thus, $n_{it}$ and $\pi_{it}$ are cointegrated if $\vartheta_i < 1$. On the contrary, (2.71) is a spurious regression if $\vartheta_i = 1$.

The tests are derived under the assumptions that $\xi_{it}$ and $\nu_{it}$ are mean zero processes. Moreover, suppose that they are cross-sectionally independent and the invariance principle applies as $T$ grows. Note that dependence across individuals is restricted to common factors[16] and, if there are multiple regressors, $\pi_{it}$ should not be cointegrated with them.

[14] Fisher (1930) said that, there is a one-to-one relation between the nominal interest rate and the inflation rate, whereas the real interest rate is constant.

[15] Possible common factors are the world real interest rate and other factors which are common among the countries that form the panel.

[16] The common factors have to fulfill the following conditions:

i. $E(u_t) = 0$, $\mathrm{Var}(u_t) < \infty$.

Westerlund (2008) based his tests on the approach suggested by Bai and Ng (2004). This approach starts with the first difference of (2.73).

$$\Delta z_{it} = \lambda_i' \Delta F_t + \Delta e_{it} \tag{2.76}$$

The common factors are estimated by applying principle components method to the OLS residuals from (2.76). To test the null hypothesis of no cointegration, a unit root test is run on the recumulated sum of the defactored and first differenced residuals. Since $\Delta z_{it}$'s are unknown, the estimates are obtained from the regression of $\Delta n_{it}$ on $\Delta \pi_{it}$. This leads to

$$\Delta \widehat{z}_{it} \quad = \quad \Delta n_{it} - \widehat{\beta}_i \Delta \pi_{it}. \tag{2.77}$$

Define $\Delta F$ and $\Delta \widehat{z}$ as $((T-1) \times m)$ and $((T-1) \times N)$ matrices of the stacked observations on $\Delta F_t$ and $\widehat{\Delta} z_{it}$, respectively. The principal components estimator of $\Delta F$ can be found by computing $\sqrt{T-1}$ times the eigenvectors of the $m$ largest eigenvalues of the $((T-1) \times (T-1))$ matrix $\Delta \widehat{z} \Delta \widehat{z}'$. Moreover, $\widehat{\lambda} = \frac{\Delta \widehat{F}' \Delta \widehat{z}}{T-1}$ is the $(m \times N)$ matrix of estimated factor loadings. As a result, the first differenced and defactored residuals used for the cointegration tests can be obtained by

$$\Delta \widehat{e}_{it} = \Delta \widehat{z}_{it} - \widehat{\lambda}_i' \Delta \widehat{F}_t. \tag{2.78}$$

Next, the residuals can be recumulated as

$$\widehat{e}_{it} = \sum_{j=2}^{t} \Delta \widehat{e}_{ij}, \tag{2.79}$$

which is a consistent estimator of $\widehat{e}_{it}$. Finally, $\widehat{e}_{it}$ can be used asymptotically to test whether $n_{it}$ and $\pi_{it}$ are cointegrated, by running a unit root test on the following regression

$$\widehat{e}_{it} = \vartheta_i \widehat{e}_{i,t-1} + \text{error}. \tag{2.80}$$

Westerlund (2008) proposed two panel cointegration tests based on the Durbin-Hausman principle. The group Durbin-Hausman statistic is

$$DH_g = \sum_{i=1}^{N} \widehat{S}_i (\widetilde{\vartheta}_i - \widehat{\vartheta}_i)^2 \sum_{t=2}^{T} \widehat{e}_{i,t-1}^2, \tag{2.81}$$

ii. $u_t$ is independent of $\xi_{it}$ and $\nu_{it}$, for all $i$ and $t$.

iii. $\frac{1}{N} \sum_{i=1}^{N} \lambda_i \lambda_i' \to \Sigma$ as $N \to \infty$, with $\Sigma > 0$.

iv. $\rho_j < 1$ for all j.

which tests $H_0 : \vartheta_i = 1$ for all $i$ vs. $H_1 : \vartheta_i < 1$ for at least one $i$. $\widehat{\vartheta}_i$ and $\widetilde{\vartheta}_i$ define the OLS and the instrumental variable (IV)[17] estimators for $\vartheta_i$, respectively. In addition, let $\widehat{S}_i = \frac{\widehat{\omega}_i^2}{(\widehat{\sigma}_i^2)^2}$, with $\widehat{\omega}_i^2$ being a consistent estimator of the long-run variance of $\xi_{it}$ and $\widehat{\sigma}_i^2$ being an estimator of the contemporaneous variance. Both variance estimates are based on the OLS residuals from (2.80).

The panel Durbin-Hausman statistic is defined as follows

$$DH_p = \widehat{S}_N(\widetilde{\vartheta} - \widehat{\vartheta})^2 \sum_{i=1}^{N} \sum_{t=2}^{T} \widehat{e}_{i,t-1}^2, \tag{2.82}$$

which tests $H_0 : \vartheta_i = \vartheta = 1$ for all $i$ vs. $H_1 : \vartheta_i = \vartheta < 1$ for all $i$. Note that the alternative hypothesis presumes a common autoregressive parameter $\vartheta_i$. In (2.82) $\widehat{\vartheta}$ and $\widetilde{\vartheta}$ denote the pooled OLS and the pooled IV estimators, respectively. Moreover, $\widehat{S}_N = \frac{\widehat{\omega}_N^2}{(\widehat{\sigma}_N^2)^2}$, with $\widehat{\omega}_N^2 = \frac{1}{N}\sum_{i=1}^{N} \widehat{\omega}_i^2$ and $\widehat{\sigma}_N^2 = \frac{1}{N}\sum_{i=1}^{N} \widehat{\sigma}_i^2$.

Both tests are asymptotically normally distributed under the null hypothesis as $N, T \to \infty$, with $N/T \to 0$. Under the alternative hypothesis the standardized Durbin-Hausman statistics diverge to positive infinity, therefore the right tail of the standard normal distribution is used for the decision of the test[18]. Note that the asymptotic distribution of the test statistics are independent of the regressors. Another advantage is that, to ensure consistency, nonstationarity is only required for the dependent variable.

If the number of the common factors $m$ is unknown, it can be estimated by minimizing an information criterion. Westerlund (2008) used the following estimator

$$\widehat{m} = \underset{0 \le m \le m_{\max}}{\arg\min} \ \ln(\widehat{\sigma}^2) + m \ln\left(\frac{NT}{N+T}\right) \frac{N+T}{NT}. \tag{2.83}$$

Here $\widehat{\sigma}^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=2}^{T}(\Delta\widehat{e}_{it})^2$ and $m_{\max}$ is a bounded integer not smaller than $m$.

With a Monte Carlo study Westerlund (2008) compared his Durbin-Hausman tests with Pedroni (2004)'s panel-$\rho$, panel-$t$, group-$\rho$ and group-$t$ tests. Although they are slightly distorted, the Durbin-Hausman tests have better size properties than the tests of Pedroni. The sizes of the latter tests

[17]For the IV estimator $\widehat{e}_{i,t-1}$ is used as an instrument. Note that the IV estimator is consistent just under the null hypothesis.

[18]If the null hypothesis is rejected, Westerlund (2008) proposed to run a panel unit root test on the dependent variable. If there is no unit root in the dependent variable, then there is no cointegration.

can even approach one if common factors are present. The powers of the Durbin-Hausman tests are also higher in comparison to the tests of Pedroni. Especially, it is advantageous to apply the Durbin-Hausman tests when $\vartheta_i$ is near unity. In case of the stationarity of the regressors and if $\beta_i$ is not predetermined, the tests of Pedroni are undersized, whereas Durbin-Hausman tests are again slightly oversized.

**Westerlund and Edgerton (2007a)**

Westerlund (2005b, 2006a) revealed by means of a Monte Carlo study that the asymptotic distribution of panel cointegration tests is a poor approximation of the empirical distribution. To overcome this problem, i.e. to improve the performance of panel cointegration tests, Westerlund and Edgerton (2007a) applied bootstrap techniques to the panel LM test of McCoskey and Kao (1998). For $i = 1, \ldots, N$; $t = 1, \ldots, T$, the random scalar $y_{it}$ and the $K$-dimensional regressor vector $x_{it}$ are defined as in (2.8) and (2.9), but the error term is now given by

$$e_{it} = \sum_{j=1}^{t} \eta_{ij} + u_{it}, \tag{2.84}$$

with $\eta_{it} \sim (0, \sigma_i^2)$ *i.i.d.* Moreover, let

$$\mathrm{w}_{it} = \sum_{j=0}^{\infty} \alpha_{ij} \xi_{i,t-j}, \tag{2.85}$$

where $\mathrm{w}_{it} = (u_{it}, \nu_{it}')'$ and $\xi_{it}$ are mean zero errors which are *i.i.d.* for all $t$. It is obvious from this structure that the test allows for heterogeneous serial correlation. In addition to this, the test accommodates for cross-sectional dependence. If the null hypothesis of cointegration panel LM test of McCoskey and Kao (1998) is applied to cross-sectionally dependent and serially correlated datasets, the test performs poorly. Thus, the null and alternative hypotheses under consideration are

$$H_0 : \sigma_i^2 = 0 \quad \text{for all } i \quad \text{vs.} \quad H_1 : \sigma_i^2 > 0 \quad \text{for at least one } i. \tag{2.86}$$

Westerlund and Edgerton (2007a) suggested using sieve bootstrap concept, in which the equilibrium errors are approximated by a finite AR process of order $p_i$ for $i = 1, \ldots, N$. Let

$$\sum_{j=0}^{\infty} \phi_{ij} \mathrm{w}_{i,t-j} = \xi_{it}, \tag{2.87}$$

which forms the basis for residual-based resampling of the bootstrap test. Therefore, the following steps should be undertaken.

a. $\phi_{ij}$ are estimated using $\widehat{\mathrm{w}}_{it} = (\widehat{e}_{it}, \nu_{it}')'$. [19]

b. The residuals are computed by

$$\widehat{\xi}_{it} = \sum_{j=0}^{p_i} \widehat{\phi}_{ij} \widehat{\mathrm{w}}_{i,t-j}. \tag{2.88}$$

Note that $\widehat{\phi}_{ij}$ are chosen with the help of the empirical Yule-Walker equations, to guarantee the invertibility of (2.88).

c. A random sample $\xi_t^*$ is drawn from the empirical distribution, which puts mass $1/T$ on each of the centered residuals $\widehat{\xi}_t - T^{-1}\sum_{j=1}^{T} \widehat{\xi}_j$, $\widehat{\xi}_t = (\widehat{\xi}_{1t}', \ldots, \widehat{\xi}_{Nt}')'$.

d. $\mathrm{w}_{it}^*$ is generated from $\xi_t^*$ using the finite version of (2.87).

e. $\mathrm{w}_{it}^*$ is partitioned into $\mathrm{w}_{it}^* = (e_{it}^*, \nu_{it}^{*\prime})'$.

f. $y_{it}^*$ and $x_{it}^*$ are generated by

$$\begin{aligned} y_{it}^* &= \widehat{\alpha}_i + x_{it}^{*\prime}\widehat{\beta}_i + e_{it}^*, & (2.89)\\ x_{it}^* &= \sum_{j=1}^{t} \Delta x_{ij}^* = \sum_{j=1}^{t} \nu_{ij}^*, & (2.90)\end{aligned}$$

in which $\widehat{\alpha}_i$ and $\widehat{\beta}_i$ are the FMOLS estimators of $\alpha_i$ and $\beta_i$, respectively.

g. The LM test statistic is computed with $y_{it}^*$ and $x_{it}^*$.

$$LM^+ = \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} \widehat{\omega}_i^{-2} S_{it}^2, \tag{2.91}$$

in which $\widehat{\omega}_i^2$ is the long-run variance of $u_{it}$ conditional on $\nu_{it}$ and $S_{it}$ is the partial sum process of $\widehat{e}_{it}$.

h. The steps from a. to g. are repeated $R$ times.

i. Finally, the 5% critical value is calculated from the lower fifth percentile of the bootstrap distribution. This empirical critical value is used instead of the asymptotic critical value.

Westerlund and Edgerton (2007a)'s Monte Carlo study showed that the bootstrap test is not oversized anymore, but it becomes undersized. Moreover, the bootstrap test is robust to cross-sectional dependence, which is not the case for the asymptotic version. However, the bootstrap test loses power against the asymptotic test.

[19] Under the null hypothesis $e_{it} = u_{it}$.

**Westerlund and Edgerton (2007b)**

Westerlund and Edgerton (2007b) proposed two panel cointegration tests for the null hypothesis of no cointegration, which accommodate for an unknown structural break both in the intercept and in the slope of the cointegrating regression. The structural breaks can be located at different dates for different cross-sections. The tests are extensions of the LM unit root tests of Schmidt and Phillips (1992), Ahn (1993) and Amsler and Lee (1995) to panel data, allowing for heteroscedastic and serially correlated errors and cross-sectional dependence.

The panel LM tests are built on the following model[20].

$$y_{it} = \delta_{0i} + \delta_{1i}t + \eta_i D_{it} + x'_{it}\beta_i + (D_{it}x_{it})'\gamma_i + z_{it}, \tag{2.92}$$

with $i = 1, \ldots, N$; $t = 1, \ldots, T$ and the $K$-dimensional regressor vector $x_{it} = x_{i,t-1} + \nu_{it}$ being $I(1)$. Moreover, $D_{it}$ represents the break dummy variable; $D_{it} = 1$ if $t > T_i^b$ and zero otherwise. $T_i^b$ represents the break date for individual $i$ such that $T_i^b = \lambda_i^b T$, with $\lambda_i^b \in [\psi, 1-\psi]$ and $\psi \in (0,1)$. To allow for cross-sectional dependence $z_{it}$ is formulated as in (2.73). Hence, the common-factors $F_{jt}$'s for $j = 1, \ldots, m$, are constructed in the same way as in (2.74). In addition to this,

$$\vartheta_i(L)\Delta e_{it} = \vartheta_i e_{i,t-1} + \xi_{it}, \tag{2.93}$$

with $\vartheta_i(L) = 1 - \sum_{j=1}^{p_i} \vartheta_{ij} L^j$ being a scalar lag polynomial. The assumptions about the error processes, i.e. $\nu_{it}$, $\xi_{it}$, $u_{it}$, and the common factors are the same as the assumptions in Westerlund (2008). In addition, some assumptions over the structural breaks are also necessary. Westerlund and Edgerton (2007b) assumed that the break points do not lie too close to the end or beginning of the sample. The second crucial assumption is that the post-break parameters of the regressors must catch up with the pre-break values as $T$ increases.

The null and alternative hypotheses are formulated as in Westerlund (2008): $H_0 : \vartheta_i = 0$ for all $i$ vs. $H_1 : \vartheta_i < 0$ for at least one $i$. If there is no cross-sectional dependence, the panel LM statistics are constructed using the following residuals.

$$\widehat{S}_{it} = y_{it} - \widehat{\delta}_{0i} - \widehat{\delta}_{1i}t - \widehat{\eta}_i D_{it} - x'_{it}\widehat{\beta}_i - (D_{it}x_{it})'\widehat{\gamma}_i, \tag{2.94}$$

[20]Based on (2.92) three different models can be evaluated. In Model 1 there is no break (i.e. $\eta_i = \gamma_i = 0$), in Model 2 there is only a break in the intercept (i.e. $\gamma_i = 0$, $\eta_i$ is unrestricted) and finally, in Model 3 there can be breaks both in the intercept and in the slope coefficients.

in which $\widehat{\delta}_{0i} = y_{i1} - \widehat{\delta}_{1i} - \widehat{\eta}_i D_{i1} - x'_{i1}\widehat{\beta}_i - (D_{i1}x_{i1})'\widehat{\gamma}_i$ is the restricted maximum-likelihood (ML) estimator of $\delta_{0i}$. The estimators for $\delta_{1i}$, $\eta_i$, $\beta_i$ and $\gamma_i$ can be found by applying OLS estimation on the first differenced regression of (2.92). Finally, the estimated regression is

$$\Delta y_{it} = \widehat{\delta}_{1i} + \widehat{\eta}_i \Delta D_{it} + \Delta x'_{it}\widehat{\beta}_i + \Delta(D_{it}x_{it})'\widehat{\gamma}_i + \Delta \widehat{z}_{it}. \tag{2.95}$$

With the help of the following equation

$$\Delta \widehat{S}_{it} = \text{constant} + \vartheta_i \widehat{S}_{i,t-1} + \text{error}, \tag{2.96}$$

the test for the null hypothesis of no cointegration is based on either the sum of the individual OLS estimates of $\vartheta_i$ from (2.96) or the sum of the individual $t$-ratios.

If there is cross-sectional dependence, then $\lambda_i$ and $\Delta F_t$ should be estimated using the same principle components procedure explained in Westerlund (2008). Next,

$$\widehat{S}_{it} = y_{it} - \widehat{\delta}_{0i} - \widehat{\delta}_{1i}t - \widehat{\eta}_i D_{it} - x'_{it}\widehat{\beta}_i - (D_{it}x_{it})'\widehat{\gamma}_i - \widehat{\lambda}'_i \widehat{F}_t, \tag{2.97}$$

is computed. $\widehat{F}_t$ and $\widehat{\lambda}_i$ are the principle components estimators of the common factors and the factor loadings, respectively. With the help of the following equation

$$\Delta \widehat{S}_{it} = \text{constant} + \vartheta_i \widehat{S}_{i,t-1} + \sum_{j=1}^{p_i} \vartheta_{ij} \Delta \widehat{S}_{i,t-j} + \text{error}, \tag{2.98}$$

the panel LM test statistics can be computed. To select the appropriate lag order $p_i$ for (2.98), either an information criterion or a sequential test based on the significance of the lag parameters $\sum_{j=1}^{p_i} \vartheta_{ij}$ can be used.

Westerlund and Edgerton (2007b) defined the panel LM test statistics as

$$\vartheta_N = \frac{1}{N}\sum_{i=1}^{N} T\widehat{\vartheta}_i S_i, \qquad \text{and} \qquad \tau_N = \frac{1}{N}\sum_{i=1}^{N} \tau_i. \tag{2.99}$$

$\widehat{\vartheta}_i$ denotes the OLS estimate of $\vartheta_i$ from (2.98) and $\tau_i$ is the $t$-ratio of $\widehat{\vartheta}_i$. Moreover, the variance ratio is defined as $S_i = \frac{\widehat{\omega}_i}{\widehat{\sigma}_i}$, in which $\hat{\omega}_i$ is the square root of kernel estimator of the long-run variance of $\Delta e_{it}$ and $\hat{\sigma}_i$ is the estimated regression standard error of (2.98). If the number of common factors $m$ is unknown, it can be estimated using the same method explained in (2.83). The unknown break point is estimated for each cross-section separately by minimizing the sum of squared residuals from the regression in (2.95).

Under the null hypothesis the standardized test statistics have a limiting standard normal distribution as $T \to \infty$ followed by $N \to \infty$. The asymptotic distribution is independent of the structural break, common factors and the number of regressors. The standardized test statistics can be formulated as

$$\frac{\sqrt{N}[\vartheta_N - E(B_i)]}{\sqrt{Var(B_i)}} \xrightarrow{w} N(0,1), \quad \frac{\sqrt{N}[\tau_N - E(C_i)]}{\sqrt{Var(C_i)}} \xrightarrow{w} N(0,1), \tag{2.100}$$

in which $E(B_i), E(C_i)$ and $Var(B_i), Var(C_i)$ are the asymptotic means and variances[21] of the underlying Brownian motion functionals. Under the alternative hypothesis the standardized statistics diverge to negative infinity, so that the left tail of the standard normal distribution is used for making the test decision. Westerlund and Edgerton (2007b) showed with a Monte Carlo study that the $\tau_N$ test has better size properties in comparison to the $\vartheta_N$ test. Depending on the adjustment method for serial correlation, the tests can be size distorted. The power increases with $T$ and when $\vartheta_i$ is different from zero. If the presence of common factors is ignored, the tests suffer from severe size distortions and low power.

### 2.1.4 Hanck Test

**Hanck (2007)**

The null hypothesis of no cointegration test of Hanck (2007) is based on the idea of combining the $p$-values from cointegration tests that are applied to each individual separately. His tests are extensions of the panel unit root tests of Maddala and Wu (1999) and Choi (2001) to panel cointegration framework, allowing for an unbalanced panel and heterogeneity in the serial correlation structure of the error terms. Moreover, this test can be applied to any cointegration test.

Hanck (2007) considered a panel consisting of $N$ cross-sections. Each cross-section has $T_i$ time observations, and a $(K+1)$-dimensional vector of variables $z_{it} = (y_{it}, x_{it}')'$, with $z_{it} = z_{i,t-1} + \mathrm{w}_{it}$. This means that all of the variables are integrated at most of order one. He also assumed that the linear error process $\mathrm{w}_{it}$ fulfills the invariance principle, and the cross-sections are independent. In addition to this, the cointegrating regression equations may have individual-specific intercepts and linear and quadratic time trends. Let $p_i$ denote the marginal significance level of the cointegration test applied to

[21]Westerlund and Edgerton (2007b) simulated the corresponding values: $E(B_i)$= -1.9675, $Var(B_i) = 0.3301$, $E(C_i) = -8.4376$ and $Var(C_i) = 25.8964$.

the cross-section $i$. The null hypothesis under consideration is

$$H_0 : \text{there is no cointegration for any } i,\ i = 1, \ldots, N, \tag{2.101}$$

and the alternative is

$$H_1 : \text{there is cointegration for at least one } i,\ i = 1, \ldots, N. \tag{2.102}$$

In addition to the assumptions above, Hanck (2007) assumed also that under the null hypothesis the chosen time series cointegration test statistic has a continuous distribution function. This ensures asymptotically a uniform $p$-value distribution of the time series test statistic under the null hypothesis, i.e. $p_i \sim U(0,1)$ for $i = 1, \ldots, N$. Thus, the test statistics he proposed are

$$\begin{aligned} P_{\chi^2} &= -2\sum_{i=1}^{N} \ln(p_i), & (2.103)\\ P_{\Phi^{-1}} &= N^{-1/2}\sum_{i=1}^{N} \Phi^{-1}(p_i), & (2.104)\\ P_t &= \sqrt{\frac{3(5N+4)}{\pi^2 N(5N+2)}}\sum_{i=1}^{N} \ln\left(\frac{p_i}{1-p_i}\right), & (2.105) \end{aligned}$$

with $\Phi^{-1}$ being the inverse of the cumulative distribution function of the standard normal distribution. These tests, to which Hanck (2007) referred as $P$ tests, are practical because they do not impose any homogeneity restrictions. Moreover, the researcher is free to choose any time series cointegration test as basis for the $P$ tests, with the restriction that it should be a cointegration test for the null hypothesis of no cointegration and the $p$-values must be available. As $T_i \to \infty$ for all $i$, under the null hypothesis the $P$ tests are asymptotically distributed as follows.

$$\begin{aligned} P_{\chi^2} &\overset{w}{\to} \chi^2_{(2N)}, & (2.106)\\ P_{\Phi^{-1}} &\overset{w}{\to} N(0,1), & (2.107)\\ P_t &\underset{\text{approx.}}{\overset{w}{\to}} t_{(5N+4)} & (2.108) \end{aligned}$$

In contrast to the tests presented so far, the asymptotic distribution is defined for finite $N$, instead of an infinite $N$.

The null hypothesis of no cointegration is rejected if $P_{\chi^2}$ exceeds the critical value from $\chi^2_{(2N)}$ distribution at $\alpha$ significance level. For $P_{\Phi^{-1}}$ and $P_t$ the null hypothesis is rejected for large negative values of the test statistics.

The crucial issue is finding the right $p$-values. To achieve this Hanck (2007) used the response surface regressions of unit root and cointegration tests proposed by MacKinnon (1996).

With a Monte Carlo study[22], Hanck (2007) revealed the properties of his $P$ tests. He used the $p$-values from the ADF test of Engle and Granger (1987) and the trace test of Johansen (1988). Briefly, the $\mathrm{P}_{\chi^2}$ test performs better than the other two tests if the $p$-values are based on the ADF test of Engle and Granger. The Johansen trace test is oversized for short and moderate panels. The powers of all tests approach unity if $T$ and $N$ increase. Note, the convergence is especially faster when the time series dimension grows.

**Hanck (2006b)**

In this study Hanck (2006b) extended the $P_{\chi^2}$ and $P_{\Phi^{-1}}$ test statistics proposed in Hanck (2007) to get robust tests in the presence of heterogeneity and cross-sectional dependence. For this purpose he used sieve bootstrap procedure with joint resampling of the residuals of different cross-sections.

Hanck (2006b) considered the following multivariate regression equation

$$y_{it} = \delta_{0i} + \delta_{1i}t + \delta_{2i}t^2 + x'_{it}\beta_i + e_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T_i, \tag{2.109}$$

in which $x_{it}$ is a $K$-dimensional vector of regressors. He derived the $p$-values from the ADF cointegration tests of Engle and Granger (1987). Hanck's $P$ tests are based on the $t$-statistic of $\rho_i - 1$ from the OLS regression

$$\Delta\widehat{e}_{it} = (\rho_i - 1)\widehat{e}_{i,t-1} + \sum_{j=1}^{p_i} \nu_j \Delta\widehat{e}_{i,t-j} + u_{it}, \tag{2.110}$$

with $\widehat{e}_{it}$ being the OLS residuals from the regression (2.109) and $u_{it}$ being white noise. In addition to the assumptions outlined above (see Hanck, 2007), the first differences of the equilibrium errors are generated by $\Delta e_{it} = \phi_i(L)u_{it}$, with $\phi_i(z) = \sum_{l=0}^{\infty} \phi_{i,l} z^l$. Hanck (2006b) defined his bootstrap algorithm as follows:

a. First the test statistics $P_{\chi^2}$ and $P_{\Phi^{-1}}$ are computed and the realizations are denoted by $\widetilde{P}_{\chi^2}$ and $\widetilde{P}_{\Phi^{-1}}$.

b. Ignoring the deterministic terms (2.109) is estimated by OLS.

$$y_{it} = \widehat{\delta}_{0i} + x'_{it}\widehat{\beta}_i + \widehat{e}_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T_i.$$

[22]Hanck (2007)'s DGP is also used by Gutierrez (2003) and Kao (1999). In their studies, they modified the DGP employed by Engle and Granger (1987), Gonzalo (1994), etc. to panel data.

c. An autoregressive process is fitted for $\Delta\widehat{e}_{it}$ and $\overline{\Delta\widehat{e}_i} = \sum_{t=2}^{T_i} \Delta\widehat{e}_{it}/(T_i-1)$ is defined to compute the empirical autocovariances of $\Delta\widehat{e}_{it}$ up to lag $q$, which are

$$\widehat{\gamma}_i(j) = \frac{\sum_{t=2}^{T_i-j}(\Delta\widehat{e}_{it} - \overline{\Delta\widehat{e}_i})(\Delta\widehat{e}_{i,t+j} - \overline{\Delta\widehat{e}_i})}{T_i - 1 - j}, \quad i, \ldots, N;\ j = 1, \ldots q.$$

The AR coefficient vector is obtained by

$$(\widehat{\phi}_{qi1}, \ldots, \widehat{\phi}_{qiq})' = \begin{pmatrix} \widehat{\gamma}_i(0) & \ldots & \widehat{\gamma}_i(q-1) \\ \vdots & \ddots & \vdots \\ \widehat{\gamma}_i(q-1) & \ldots & \widehat{\gamma}_i(0) \end{pmatrix}^{-1} \widehat{\gamma}_i;\ i = 1, \ldots, N,$$

with $\widehat{\gamma}_i = (\widehat{\gamma}_i(1), \ldots, \widehat{\gamma}_i(q))'$. Any information criterion can be used for selecting the lag order $q$. Moreover, $q$ can vary over individuals.

d. Then, the residuals are

$$\widehat{u}_{qit} = \Delta\widehat{e}_{it} - \sum_{l=1}^{q} \widehat{\phi}_{qil} \Delta\widehat{e}_{i,t-l}, \quad i = 1, \ldots, N; \quad t = q+2, \ldots, T_i,$$

and $\widehat{u}_{qit}$ can be centered by

$$\widetilde{u}_{qit} = \widehat{u}_{qit} - \frac{\sum_{g=q+2}^{T_i} \widehat{u}_{qig}}{T_i - q - 1}, \quad i = 1, \ldots, N; \quad t = q+2, \ldots, T_i.$$

e. $u^*_{qit}$ is obtained by resampling non-parametrically from $\widetilde{u}_{qit}$. Note that resampling should be done jointly to maintain the cross-sectional dependence structure.

$$\widetilde{u}_{q.t} = (\widetilde{u}_{q1t}, \ldots, \widetilde{u}_{qNt}), \quad t = q+2, \ldots, T_i.$$

f. The bootstrap samples can be established by

$$\Delta e^*_{qit} = \sum_{l=1}^{q} \widehat{\phi}_{qil} \Delta e^*_{qi,t-l} + u^*_{qit}, \quad i = 1, \ldots, N; \quad t = q+2, \ldots, T_i.$$

g. While generating the artificial data, the null hypothesis of no cointegration is imposed by integrating $\Delta e^*_{it}$ to get $e^*_{it}$. Let

$$y^*_{qit} = \widehat{\alpha}_i + x'_{it}\widehat{\beta}_i + e^*_{qit}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T_i.$$

h. The $P$ tests are computed with artificial data $(y^*_{qit}, x'_{it})'$. The test statistics are symbolized by $P^{b*}_{\chi^2}$ or $P^{b*}_{\Phi^{-1}}$.

i. The steps from a. to h. are replicated $R$ times.

j. The null hypothesis of no cointegration is rejected if

$$\frac{\sum_{b=1}^{R} \mathbf{1}\{P^{b*}_{\chi^2} > \widetilde{P}_{\chi^2}\}}{R} < \alpha \quad \text{or} \quad \frac{\sum_{b=1}^{R} \mathbf{1}\{P^{b*}_{\Phi^{-1}} < \widetilde{P}_{\Phi-1}\}}{R} < \alpha.$$

Here $\mathbf{1}\{\cdot\}$ is the indicator function.

Hanck (2006b) did not formally prove consistency of the bootstrap procedure. Instead, he referred to the comment of Chang et al. (2006) that their proof for consistency can be generalized also for panel cointegration tests. Note, one should be careful with the choice of the lag order $q$ because the size and power of the test can be affected by the selection of the correct lag order. However, this is also a problem in semi-parametric and parametric panel cointegration tests.

The Monte Carlo study shows that the $P$ tests are oversized if there is cross-sectional dependence. On the contrary, the bootstrap $P$ tests have better size properties than the simple $P$ tests, which depend on the selection of the appropriate lag order. The power increases with $T$, but the gain in power is less in case of strong cross-sectional dependence.

### 2.1.5 Banerjee and Carrion-i-Silvestre Test

Banerjee and Carrion-i- Silvestre (2006) extended the parametric panel-$\rho$ ($Z_{\widehat{\rho}_{NT}}$) and the panel-$t$ ($Z_{t_{NT}}$) test statistics of Pedroni (1999, 2004) to tests that allow for structural breaks in the deterministic terms and/or in the cointegrating vector. In addition to this, they also generalized the tests of Pedroni (2004) to accommodate for cross-sectional dependence by applying the common factor modelling of Bai and Ng (2004). With a Monte Carlo study Banerjee and Carrion-i- Silvestre (2006) showed that the tests of Pedroni (2004) lose power when the structural break is omitted. The loss in power is more severe if the break point is in the middle or at the end of the time series.

#### Cross-sectional independence

Banerjee and Carrion-i- Silvestre (2006) considered six different models under the assumption that there is no cross-sectional dependence.

- Model 1: Only a shift in the intercept: $y_{it} = \alpha_{1i} + \alpha_{2i}DU_{it} + x'_{it}\beta_i + z_{it}$,
- Model 2: Time trend and a shift in the intercept: $y_{it} = \alpha_{1i} + \alpha_{2i}DU_{it} + \theta_{1i}t + x'_{it}\beta_i + z_{it}$,
- Model 3: Shifts in the intercept and in the time trend: $y_{it} = \alpha_{1i} + \alpha_{2i}DU_{it} + \theta_{1i}t + \theta_{2i}DT^*_{it} + x'_{it}\beta_i + z_{it}$,
- Model 4: Shifts in the intercept and in the cointegrating vector: $y_{it} = \alpha_{1i} + \alpha_{2i}DU_{it} + x'_{it}\beta_{it} + z_{it}$,
- Model 5: Time trend and shifts in the intercept and in the cointegrating vector: $y_{it} = \alpha_{1i} + \alpha_{2i}DU_{it} + \theta_{1i}t + x'_{it}\beta_{it} + z_{it}$,
- Model 6: Shifts in the intercept, in the time trend and in the cointegrating vector: $y_{it} = \alpha_{1i} + \alpha_{2i}DU_{it} + \theta_{1i}t + \theta_{2i}DT^*_{it} + x'_{it}\beta_{it} + z_{it}$,

Again, $i = 1, \ldots, N$; $t = 1, \ldots, T$ and $x_{it} = x_{i,t-1} + \nu_{it}$. The dummy variables are

$$DU_{it} = \begin{cases} 0, & \text{if} \quad t \leq T_i^b \\ 1, & \text{if} \quad t > T_i^b \end{cases}, \quad \text{and} \quad DT^*_{it} = \begin{cases} 0, & \text{if} \quad t \leq T_i^b \\ (t - T_i^b), & \text{if} \quad t > T_i^b \end{cases},$$

with $T_i^b = \lambda_i T$ representing the break date for $i$ and $\lambda_i \in (0,1)$. The random process $\mathrm{w}_{it} = (z_{it}, \nu'_{it})'$ satisfies the invariance principle, and $z_{it}$ and $\nu_{it}$ are independent of each other. Following the procedure in Gregory and Hansen (1996a), to compute the generalized Pedroni tests, first the models described above are estimated by OLS. Next, the parameters of an ADF regression of the OLS residuals[23] $\widehat{z}_{it}(\lambda_i)$ is estimated. The break date $T_i^b$ can be estimated by minimizing the individual ADF statistics

$$\widehat{T}_i^b = \arg \min_{\lambda_i \in (0,1)} t_{\widehat{\rho}_i}(\lambda_i), \quad \widehat{T}_i^b = \arg \min_{\lambda_i \in (0,1)} T\widehat{\rho}_i(\lambda_i), \quad \text{for all } i = 1, \ldots, N, \tag{2.111}$$

in which $t_{\widehat{\rho}_i}(\lambda_i)$ is the $t$-ratio and $T\widehat{\rho}_i(\lambda_i)$ denotes the normalized bias[24].

Banerjee and Carrion-i- Silvestre (2006) defined the generalized panel-$\rho$ and panel-$t$ statistics of Pedroni as

$$N^{-1/2} Z_{\widehat{\rho}_{NT}}(\widehat{\lambda}) = N^{-1/2} \sum_{i=1}^{N} T\widehat{\rho}_i(\widehat{\lambda}_i), \tag{2.112}$$

$$N^{-1/2} Z_{t_{NT}}(\widehat{\lambda}) = N^{-1/2} \sum_{i=1}^{N} t_{\widehat{\rho}_i}(\widehat{\lambda}_i). \tag{2.113}$$

---

[23] $\Delta\widehat{z}_{it}(\lambda_i) = \rho_i \widehat{z}_{i,t-1}(\lambda_i) + \phi_{i1}\Delta\widehat{z}_{i,t-1}(\lambda_i) + \ldots + \phi_{ip}\Delta\widehat{z}_{i,t-p}(\lambda_i) + \varepsilon_{it}$.

[24] Based on the ADF regression of $\Delta\widehat{z}_{it}(\lambda_i)$, the normalized bias $T\widehat{\rho}_i(\lambda_i)$ can be computed as $T\widehat{\rho}_i(\lambda_i) = T\widehat{\rho}_i(1 - \widehat{\phi}_{i1} - \ldots, \widehat{\phi}_{ip})^{-1}$.

$\widehat{\lambda} = (\widehat{\lambda}_1, \widehat{\lambda}_2, \ldots, \widehat{\lambda}_N)'$ is the estimated vector of break fractions.

Under the null hypothesis the standardized $Z_{\widehat{\rho}_{NT}}$ and $Z_{t_{NT}}$ statistics converge to standard normal distribution as $T \to \infty$ followed by $N \to \infty$. Moreover, under the alternative hypothesis they diverge to negative infinity. Since the asymptotic moments can be a poor approximation in finite samples, Banerjee and Carrion-i- Silvestre (2006) approximated the finite sample moments with response surfaces, which are functions of $T$ and the number of regressors.

Assuming that the break point is unknown and the cross-sections are independent, Banerjee and Carrion-i- Silvestre (2006) showed in a Monte Carlo study that the sizes of the $Z_{\widehat{\rho}_{NT}}$ and $Z_{t_{NT}}$ tests are around the 5% nominal significance level, for all considered $T$ and $N$ values. In terms of power, $Z_{t_{NT}}$ outperforms $Z_{\widehat{\rho}_{NT}}$. The power of $Z_{\widehat{\rho}_{NT}}$ is only higher if the estimated model coincides with the true DGP. In addition, the power of $Z_{t_{NT}}$ is not affected by deviations of the estimated model from the true DGP.

### Cross-sectional dependence

Banerjee and Carrion-i- Silvestre (2006) reconsidered Model 1 to Model 6 to define the panel cointegration test which accounts for cross-sectional dependence. In addition to equations for $y_{it}$ and $x_{it}$ explained in the previous subsection, they defined

$$
\begin{aligned}
z_{it} &= \pi_i' F_t + e_{it}, && (2.114)\\
F_{jt} &= \psi_j F_{j,t-1} + u_{jt}, && (2.115)\\
e_{it} &= \phi_i e_{i,t-1} + \xi_{it}, && (2.116)
\end{aligned}
$$

in which $F_t = (F_{1t}, \ldots, F_{mt})'$ is an $(m \times 1)$ vector of common factors and $\pi_i$ is an $(m \times 1)$ vector of factor loadings. The errors $\nu_{it}$, $e_{it}$ and $u_t = (u_{1t}, \ldots, u_{mt})'$ are mean zero *i.i.d.* stationary processes. The regressors are assumed to be strictly exogenous. Moreover, $\xi_{it}$, $u_t$ and $\pi_i$ are mutually independent.

Similar to the procedure defined in Bai and Ng (2004), which is also implemented by Westerlund and Edgerton (2007b), the common factors and the factor loadings are estimated with the principal components method[25] by taking the first differences of $y_{it}$ and $z_{it}$. Using the orthogonal projection matrix $M_i = I - \Delta x_i^d (\Delta x_i^{d\prime} \Delta x_i^d)^{-1} \Delta x_i^{d\prime}$, in which $\Delta x_i^d$ captures $\Delta x_i$ and the differenced forms of the deterministic terms, the first differenced $y_{it}$ equation turns into

$$M_i \Delta y_{it} = M_i \pi_i' \Delta F_t + M_i \Delta e_{it}. \tag{2.117}$$

[25] For further details on principal components estimator please see the subsection Westerlund and Edgerton (2007b) in Section 2.1.3.

With the principal components estimates of $F_t$ and $\pi_i$, which can be represented as $\widetilde{F}_t$ and $\widetilde{\pi}_i$, respectively, the estimated residuals are defined as

$$M_i \Delta \widetilde{e}_{it} = M_i \Delta y_{it} - M_i \widetilde{\pi}_i' \Delta \widetilde{F}_t. \tag{2.118}$$

Hence, the disturbance terms are derived by $\widetilde{e}_{it} = \sum_{j=2}^{t} (y_{ij}^{*} - \widetilde{f}_j' \widetilde{\pi}_i)$, with $y_{it}^{*} = M_i \Delta y_{it}$ and $\widetilde{f} = (\widetilde{f}_2, \ldots, \widetilde{f}_T) = M_i \Delta \widetilde{F}$.

The null hypothesis of no cointegration is tested again by implementing a unit root test on the ADF regression of $\widetilde{e}_{it}$

$$\Delta \widetilde{e}_{it}(\widehat{\lambda}_i) = \alpha_{i0} \widetilde{e}_{i,t-1}(\widehat{\lambda}_i) + \alpha_{i1} \Delta \widetilde{e}_{i,t-1}(\widehat{\lambda}_i) + \ldots + \alpha_{ip} \Delta \widetilde{e}_{i,t-p}(\widehat{\lambda}_i) + \varepsilon_{it}. \tag{2.119}$$

As a result the null hypothesis of no cointegration can be formulated as

$$H_0 : \alpha_{i0} = 0, \text{ for all } i. \tag{2.120}$$

Based on (2.119) the $t$-statistic for $\alpha_{i0} = 0$ can be computed, which is denoted by $t_{\widetilde{e}_i}^{j}(\lambda_i)$ for $j = 1, \ldots, 6$, and $j$ defines the six different models.

As $T$ and $N$ approach infinity sequentially, the limiting distributions of the ADF statistics are functionals of Brownian motion and they do not depend on the regressors. Moreover, the breaks do not influence the limiting distribution, unless there is a break in the trend. Then, the limiting distribution of the statistics depends on the number and location of the breaks. To overcome this problem Banerjee and Carrion-i- Silvestre (2006) assumed that the break dates are homogeneous for all $i$ (i.e. $\lambda_i = \lambda$ for $i = 1, \ldots, N$). In this way as $M_i$ is independent of $i$ asymptotically, the tests will be correctly sized.

The panel cointegration statistics for the models without any trend break can be formulated as

$$Z_{t_{NT}}^{\widetilde{e}^j}(\lambda) = \sum_{i=1}^{N} t_{\widetilde{e}_i}^{j}(\lambda_i), \quad \text{for} \quad j = 1, 2, 4, 5, \tag{2.121}$$

with $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_i, \ldots, \lambda_N)'$ and $j$ denotes the considered model. For the models with trend change the panel statistic is

$$Z_{t_{NT}}^{\widetilde{e}^\gamma}(\lambda) = \sum_{i=1}^{N} t_{\widetilde{e}_i}^{\gamma}(\lambda), \quad \text{for} \quad \gamma = 3, 6, \tag{2.122}$$

in which the break date is common for all $i$, i.e. $\lambda_1 = \lambda_2 = \ldots = \lambda_N$, and $\gamma$ denotes the model under consideration.

If the panel statistics are standardized with the appropriate asymptotic moments, under the null hypothesis the statistics have a standard normal distribution as $T \to \infty$ and $N \to \infty$ sequentially. Under the alternative hypothesis they diverge to negative infinity.

If there is only one common factor, Banerjee and Carrion-i- Silvestre (2006) suggested to employ the DF test to the detrended data $\widetilde{F}_t$ obtained from principle components analysis. In the presence of more than one factor they applied the MQ statistic of Bai and Ng (2004) to determine the number of stochastic trends. However, the simulation study demonstrates that the tests have poor properties.

Heterogeneous unknown break dates are estimated by minimizing the sum of squared residuals over all possible break dates using $y_{it}$ in first differences. Having estimated the break dates, the unknown parameters are estimated and the standardized panel tests statistics can be computed. If the unknown break date is homogeneous, it is estimated by computing $\widetilde{Z^e_{t_{NT}}}(\lambda)$ for each possible date under the assumption that the break point is common for all $i$. Thus, the estimated break date $\widehat{T}^b$ is the argument that minimizes the sequence of standardized $\widetilde{Z^e_{t_{NT}}}(\lambda)$ statistics.

Based on a DGP which allows for cross-sectional dependence, Banerjee and Carrion-i- Silvestre (2006) demonstrated that $\widetilde{Z^e_{t_{NT}}}$ has good size and power properties in finite samples.

### 2.1.6 Gengenbach, Palm and Urbain Test

Using already existing tools, Gengenbach et al. (2006) proposed panel cointegration test for the null hypothesis of no cointegration, which allow for cross-sectional dependence in the panel[26]. Gengenbach et al. (2006) assumed that the nonstationarity of the variables originates either from the common or the idiosyncratic stochastic trends. Therefore, they suggested to defactor the model before testing. Gengenbach et al. (2006) considered the following model[27]

$$
\begin{aligned}
z_{it} &= \Lambda_i F_t + E_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, && (2.123)\\
F_t &= F_{t-1} + f_t, && (2.124)\\
f_t &= \Phi(L)\eta_t, \quad \Phi(L) = \sum_{j=0}^{\infty} \phi_j L^j, && (2.125)
\end{aligned}
$$

in which $z_{it} = (y_{it}, x'_{it})'$, $F_t$ is an $(m \times 1)$ vector of common $I(1)$ factors, and $m$-dimensional column vector $\eta_t$ is $\eta_t \sim (0, I_m)$ *i.i.d.* with finite fourth

[26]see Bai and Ng, 2004 for a detailed description of the cross-sectional dependence modelling.

[27]Note that individual specific deterministic terms can be added to the model. But, for brevity Gengenbach et al. (2006) included no deterministic terms without loss of generality.

moments. The block-diagonal $((K+1) \times m)$ matrix of factor loadings $\Lambda_i$ is assumed to have full rank. The block diagonality of $\Lambda_i$ ensures that $y_{it}$ and $x_{it}$ are not cointegrated if nonstationarity is caused only by the common factors. Gengenbach et al. (2006) partitioned the matrices $F_t$, $\Lambda_i$ and $E_{it}$ according to the partition of $z_{it}$, so that (2.123) can be rewritten as

$$z_{it} = \begin{bmatrix} \lambda'_{1i} & 0 \\ 0 & \lambda'_{2i} \end{bmatrix} \begin{bmatrix} F_t^Y \\ F_t^X \end{bmatrix} + \begin{bmatrix} E_{it}^Y \\ E_{it}^X \end{bmatrix}. \qquad (2.126)$$

$F_t^Y$ and $\lambda_{1i}$ are $(m_Y \times 1)$ vectors, $F_t^X$ and $\lambda_{2i}$ are $(m_X \times 1)$ vectors, $E_{it}^Y$ is a scalar and $E_{it}^X$ is a $(K \times 1)$ vector. They considered two different cases. In the first case they assumed the stationarity of the idiosyncratic component $E_{it} = e_{it}$, and in the second case they assumed that $E_{it}$ is nonstationary, i.e $E_{it} = E_{i,t-1} + e_{it}$, $e_{it} = \left(\sum_{j=0}^{\infty} \Gamma_{ij} L^j\right) \varepsilon_{it}$ and $\varepsilon_{it} \sim (0, \Sigma_i)$ *i.i.d.* For each $i$, $\varepsilon_{it}$ has finite eight moments and $\varepsilon_{it}$ and $\varepsilon_{js}$ are independent for any $t$, $s$ and $i \neq j$.

Moreover, it is assumed that there is no cointegration between the common factors and they form an $m$-dimensional $I(1)$ process. To estimate the model, the factor loadings should be identifiable, and the invariance principle should hold for $F_t$ and $S_{it} = \sum_{s=1}^{t} e_{is}$. In addition, $\eta_t$, $\varepsilon_{it}$ and $\Lambda_i$ are mutually independent. Finally, if $E_{it} = E_{i,t-1} + e_{it}$, then the idiosyncratic components are not cointegrated along the cross-section.

Gengenbach et al. (2006) suggested a two-step procedure for testing the null of no cointegration in panels with common factors. With their approach three different hypothesis can be tested: (a) the null hypothesis of no cointegration in idiosyncratic component, (b) the null hypothesis of no cointegration in common factors, and finally (c) the null hypothesis of no panel cointegration. In the first step of the estimation procedure, the common factors are estimated using the principal component analysis proposed by Bai and Ng (2004). Additionally, unit root tests are run on the factors and the idiosyncratic components. In the second step, procedure is split up into two different cases. If there is cross-sectional cointegration, i.e. the common factors are $I(1)$ and the idiosyncratic components are $I(0)$, cointegration between $y_{it}$ and $x_{it}$ exists if and only if the common factors of $y_{it}$ and $x_{it}$ cointegrate[28].

If both the common factors and the idiosyncratic components are nonstationary, first the null hypothesis of no cointegration between the estimated common factors of $y_{it}$ and $x_{it}$ is tested, and then the null hypothesis of no

[28] Cointegration between $y_{it}$ and $x_{it}$ can be detected by testing for the null hypothesis of no cointegration between the common factors using Johansen's LR tests.

cointegration between the defactored $y_{it}$ and $x_{it}$ is tested using the panel cointegration tests of Pedroni (1999, 2004).

The null hypothesis of no panel cointegration is rejected if the hypotheses (a) and (b) outlined above are rejected and the restrictions between the cointegrating vector parameters cannot be rejected. Please note that this procedure requires large $T$ and $N$ and only deals with a single cointegrating relation. The Monte Carlo study of Gengenbach et al. (2006) shows that the proposed test procedure has low power and size distortions.

### 2.1.7 Gutierrez Test

Gutierrez (2008) suggested three different panel tests for the null hypothesis of no cointegration, which are extensions of Gregory and Hansen (1996a,b) tests. The new tests only allow for one unknown structural break in the slope, intercept and/or time trend for at least one cross-section.

Although the tests of Westerlund (2006c) are also based on Gregory and Hansen (1996a), the tests of Gutierrez (2008) are quite different in certain aspects. The tests of Westerlund (2006c) only allow for one structural break in the intercept, and the first two moments of the asymptotic test statistics are necessary for the applicability of the tests. By following the procedure of Maddala and Wu (1999), Gutierrez (2008) combined the $p$-values of the test statistics, which are computed for each cross-section separately, and there is no need to simulate the asymptotic moments. In this aspect, the tests of Gutierrez (2008) resemble the panel cointegration tests of Hanck (2007).

Depending on, in which component of the process the break occurs, the cointegrating regression model with a single regressor[29] takes in five different forms:

- Model 1: Only a shift in the intercept: $y_{it} = \alpha_{1i} + \alpha_{2i} DU_{it} + x_{it}\beta_{1i} + e_{it}$.
- Model 2: Time trend and a shift in the intercept: $y_{it} = \alpha_{1i} + \alpha_{2i} DU_{it} + \theta_{1i} t + x_{it}\beta_{1i} + e_{it}$.
- Model 3: Shifts in the intercept and in the time trend: $y_{it} = \alpha_{1i} + \alpha_{2i} DU_{it} + \theta_{1i} t + \theta_{2i} DU_{it} t + x_{it}\beta_{1i} + e_{it}$.
- Model 4: Shifts in the intercept and in the cointegrating vector: $y_{it} = \alpha_{1i} + \alpha_{2i} DU_{it} + x_{it}\beta_{1i} + x_{it}\beta_{2i} DU_{it} + e_{it}$.
- Model 5: Shifts in the time trend and in the cointegrating vector: $y_{it} = \alpha_{1i} + \theta_{1i} t + \theta_{2i} DU_{it} t + x_{it}\beta_{1i} + x_{it}\beta_{2i} DU_{it} + e_{it}$.

[29] Note that the procedure can be extended to the case with more than one regressors.

Throughout the specifications, $\beta_{1i}$ is the slope parameter, $e_{it}$ is a stationary error process, and both, $y_{it}$ and $x_{it} = x_{i,t-1} + \nu_{it}$ are $I(1)$, for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Moreover, the shift dummy is defined as

$$DU_{it} = \begin{cases} 0, & \text{if} \quad t \leq T_i^b \\ 1, & \text{if} \quad t > T_i^b \end{cases} \quad \text{for} \quad i = 1, \ldots, N, \tag{2.127}$$

in which $T_i^b = \lambda_i T$ denotes the date of the break and $\lambda_i \in (0, 1)$ is a fixed fraction of $T$. Like the other residual-based test statistics discussed above, $\mathrm{w}_{it} = (e_{it}, \nu_{it}')'$ satisfies the invariance principle, and $\mathrm{w}_{it}$ is independent over the cross-section dimension. Under the null hypothesis $e_{it}$ is an $I(1)$ process and no shift occurs, i.e. $DU_{it} = 0$, whereas under the alternative hypothesis $e_{it}$ is $I(0)$ and $DU_{it} \neq 0$ at least for one $i$.

In line with Gregory and Hansen (1996a,b), Gutierrez (2008) first computed $Z_\alpha$, $Z_t$ and the ADF statistics of Phillips and Ouliaris (1990) for each $i$. The smallest values of these statistics, which are obtained for different $\lambda_i$, are used as final test statistic. For $i = 1, \ldots, N$ and the five different models discussed above, $j = 1, \ldots, 5$, the test statistics can be formulated as follows.

$$ADF_{ji}^* = \arg \min_{\lambda_i \in (0,1)} ADF_{ji}(\lambda_i), \tag{2.128}$$

$$Z_{\alpha,ji}^* = \arg \min_{\lambda_i \in (0,1)} Z_{\alpha,ji}(\lambda_i), \tag{2.129}$$

$$Z_{t,ji}^* = \arg \min_{\lambda_i \in (0,1)} Z_{t,ji}(\lambda_i). \tag{2.130}$$

Please note that according to Gregory and Hansen (1996a,b) the distributions of the test statistics do not depend on the parameter $\lambda_i$.

Next, by applying the procedure of Maddala and Wu (1999), the asymptotic-$p$ values $p_i$, $i = 1 \ldots, N$, of the three statistics discussed above are computed. Thus, the three panel test statistics are formulated as follows:

$$P_{\chi^2} = -N^{-1/2} \sum_{i=1}^{N} (\ln(p_i) + 1), \tag{2.131}$$

which is a modified version of Fisher (1932)'s inverse $\chi^2$-test,

$$Z = N^{-1/2} \sum_{i=1}^{N} \Phi^{-1}(p_i), \tag{2.132}$$

is the inverse normal test, in which $\Phi^{-1}$ denotes the inverse of the standard normal cumulative distribution function, and

$$Z_L = (\pi^2 N/3)^{-1/2} \sum_{i=1}^{N} \ln \left( \frac{p_i}{1 - p_i} \right), \tag{2.133}$$

is a modified version of a logit test. Under the null hypothesis as $T \to \infty$ and $N \to \infty$, the test statistics (2.131)-(2.133) have standard normal limiting distributions. Moreover, under the alternative $P_{\chi^2}$ diverges to positive infinity, so the right tail of the standard normal distribution is considered to reject the null hypothesis. On the contrary, both $Z$ and $Z_L$ diverge to negative infinity, so the left tail of the standard normal distribution is used to reject the null hypothesis.

### 2.1.8 Other Residual Based Panel Cointegration Tests

Fachin (2007) proposed bootstrap panel cointegration tests appropriate for small samples, which allows for short- and long-run cross-sectional dependence. The tests are implemented by using the continuous-path block bootstrap technique developed by Paparoditis and Politis (2001). Fachin (2007) focused on the group-$t$ statistics of Pedroni (2004) and the median of the individual ADF statistic. Cross-sectional dependence is attained by a common factor modelling. In a Monte Carlo study Fachin (2007) showed that the bootstrap tests are more robust to short-run and long-run cross-sectional dependence compared to the asymptotic version of the tests.

Di Iorio and Fachin (2007) developed bootstrap panel cointegration tests which use again the continuous-path block bootstrap method. These panel cointegration tests accommodate not only cross-sectional dependence, but also heterogeneous breaks in the slope coefficients of the regression equation. The cross-sectional dependence is modelled by assuming nonstationary common factors. Di Iorio and Fachin (2007) based their tests on the mean and median of the individual ADF test statistic proposed by Gregory and Hansen (1996a). The tests allow for breaks at unknown dates and are developed for the null hypothesis of no cointegration. The Monte Carlo study reveals that the sizes of the bootstrap tests converge to the nominal size level and the powers are quite high. The tests are undersized when the breaks are at the end of the sample or whenever the breaks take place in different intervals of the sample for different cross-sections.

Tam (2007) dealt with panel cointegration testing in the presence of structural breaks. First, she compared the Gregory and Hansen (1996a) type tests of Banerjee and Carrion-i- Silvestre (2006) and the LM type tests of Westerlund and Edgerton (2007b) using a Monte Carlo study. She considered a DGP with linear trend term and took different cases into account in which the break can be either in the intercept, and/or in the trend parameter and/or in the cointegrating vector. For the Monte Carlo study, Tam (2007) modified the calculation of the existing test statistics. She estimated the adjustment factor for the $T\widehat{\rho}_i(\lambda_i)$ of Banerjee and Carrion-i- Silvestre (2006) tests using the same method as Westerlund and Edgerton (2007b), i.e. with $S_i = \frac{\widehat{\omega}_i}{\widehat{\sigma}_i}$, instead of estimating it with $(1 - \widehat{\phi}_{i1} - \ldots, \widehat{\phi}_{ip})^{-1}$. Tam (2007) also modified the break date selection method for both tests. To select the break date for the Banerjee and Carrion-i- Silvestre (2006) tests, instead of minimizing the individual test statistics, she proposed to minimize the sum of squared residuals of the ADF regression used to calculate the test statistics. For the tests of Westerlund and Edgerton (2007b), Tam (2007) chose to minimize the sum of squared residuals of the first differenced regression of $y_{it}$. Instead of the asymptotic moments she used the finite sample moments, which depend on $T$ and the lag order of the ADF regression. Simulation results demonstrate that the modified tests perform better than the original tests. Hence, the tests of the Banerjee and Carrion-i- Silvestre (2006), which do not allow for breaks under the null hypothesis, are less powerful than the tests of Westerlund and Edgerton (2007b), which allow for breaks both, under the null and alternative hypotheses. Furthermore, the results also show that the modified break date selection method is more accurate than the selection methods originally used in Banerjee and Carrion-i- Silvestre (2006) and Westerlund and Edgerton (2007b). Moreover, Tam (2007) extended the LM tests of Westerlund and Edgerton (2007b) to the cases with breaks in the trend term, in addition to the breaks in the intercept and/or cointegrating vector.

Under the assumption of cross-sectional dependence, Tam (2007) proposed another break date selection method based on the minimization of the aggregate sum of squared residuals of the regression under the null hypothesis. Her Monte Carlo study reveals that this method selects the true break date more accurately than the methods used in Banerjee and Carrion-i-Silvestre (2006) and Westerlund and Edgerton (2007b).

## 2.2 Maximum-Likelihood-Based Tests

### 2.2.1 Larsson, Lyhagen and Löthgren (LLL) Test

Larsson et al. (2001) presented a maximum-likelihood-based panel test for the cointegrating rank in heterogeneous panels. They proposed a standardized LR-bar test based on the mean of the individual rank trace statistic of Johansen (1995).

Larsson et al. (2001) considered the following heterogeneous panel VAR model, for the $K$-dimensional process $y_{it}$

$$y_{it} = \sum_{j=1}^{p_i} A_{ij} y_{i,t-j} + \varepsilon_{it}, \quad i = 1, \ldots, N; \; t = 1, \ldots, T, \tag{2.134}$$

which allows neither an intercept nor a time trend in the VAR model. The error process $\varepsilon_{it}$ is assumed to be Gaussian white noise with a nonsingular covariance matrix, i.e. $\varepsilon_{it} \sim N_K(0, \Omega_i)$, and the initial conditions $y_{i,-p_i+1}, \ldots, y_{i0}$ are fixed. Consider the following error correction representation of (2.134).

$$\Delta y_{it} = \Pi_i y_{i,t-1} + \sum_{j=1}^{p_i-1} \Gamma_{ij} \Delta y_{i,t-j} + \varepsilon_{it}, \quad i = 1, \ldots, N;\ t = 1, \ldots, T, \quad (2.135)$$

with the $(K \times K)$ matrix $\Pi_i = -(I_K - A_{i1} - \ldots - A_{i,p_i})$, $i = 1, \ldots, N$. $\Pi_i$ can be decomposed to $\Pi_i = \alpha_i \beta_i'$, in which $\alpha_i$ and $\beta_i$ are $(K \times r_i)$ matrices with full column rank. The short-run parameter matrices $\Gamma_{ij}$ are defined by $\Gamma_{ij} = -(A_{i,j+1} + \ldots + A_{i,p_i})$ for $j = 1, \ldots, p_i - 1$.

Larsson et al. (2001) considered the null hypothesis that in all of the $N$ cross-sections there are at most $r$ cointegrating relations among the $K$ variables. Thus, the null hypothesis for the panel cointegration test can be expressed as

$$H_0 : \text{rank}(\Pi_i) = r_i \leq r, \qquad \text{for all } i = 1, \ldots, N, \quad (2.136)$$

which is tested against the alternative

$$H_1 : \text{rank}(\Pi_i) = K, \qquad \text{for all } i = 1, \ldots, N. \quad (2.137)$$

Note that the hypotheses imply a sequential testing procedure. Following the testing procedure of Johansen (1988), first $H_0 : \text{rank}(\Pi_i) = r_i \leq 0$ is tested for all $i$, and if $\text{rank}(\Pi_i) = r_i \leq 0$ is rejected $H_0 : \text{rank}(\Pi_i) = r_i \leq 1$ is tested. The procedure continues until the null hypothesis is not rejected or $H_0 : \text{rank}(\Pi_i) = r_i \leq K - 1$ is rejected.

The standardized LR-bar statistic for the panel cointegrating rank test is defined by

$$\Upsilon_{\overline{\text{LR}}}\{H(r)|H(K)\} = \frac{\sqrt{N}\left[\frac{1}{N}\sum_{i=1}^{N}\left(-T\sum_{j=r+1}^{K}\ln(1-\widehat{\lambda}_{ij})\right) - E(Z_d)\right]}{\sqrt{Var(Z_d)}}.$$

Here $\widehat{\lambda}_{i1} \geq \ldots \geq \widehat{\lambda}_{iK}$ are the ordered generalized eigenvalues for cross-section $i$ which are obtained by the eigenvalue problem defined in Johansen (1995). $E(Z_d)$ is the mean and $Var(Z_d)$ is the variance of the asymptotic trace statistic

$$Z_d = \text{tr}\left[\int_0^1 \mathrm{d}W(s)W(s)'\left(\int_0^1 W(s)W(s)'\right)^{-1}\int_0^1 W(s)\mathrm{d}W(s)'\right]. \quad (2.138)$$

$W(s)$ is a $d = K - r$ dimensional standard Brownian motion with identity covariance matrix. Larsson et al. (2001) simulated the mean and variance of the asymptotic trace statistic $Z_d$ for different $d$ values using the simulation procedure described in Johansen (1995)[30].

The test is applied under the assumption that the variables are integrated at most of order one. Moreover, the test does not allow for cross-sectional dependence. This strong assumption may cause size distortions if the assumption is violated. To establish the asymptotic distribution of the standardized LR-bar statistic with the central limit theorem, the first two moments of the asymptotic trace statistic must exist and be finite. However, in Chapter 4 it will be demonstrated that the proof related to the existence of the first two moments of the asymptotic trace statistic in Larsson et al. (2001) is not correct[31].

Hence, when the first two moments of asymptotic trace statistic are finite, then under the null hypothesis the standardized LR-bar statistic is standard normally distributed as $N$ and $T \to \infty$ in such a way that $\sqrt{N}/T \to 0$. To accomplish this joint convergence, the short-run dynamics are allowed to vary over cross-sections and the long-run dynamics are assumed to be the same for all cross-sections, i.e. $\Pi = \alpha\beta'$ for all $i$.[32]

The panel cointegrating rank test of Larsson et al. (2001) is a one-sided test and the null hypothesis is rejected for all $i$ if the standardized LR-bar statistic is larger than the $(1-\alpha)$ standard normal quantile, with $\alpha$ as the significance level of the test.

In their simulation study, Larsson et al. (2001) found out that for panels with small $T$, the standardized LR-bar test is oversized and has low power. Moreover, the size and the power of the test increase for large $T$, but the size does not approach the nominal significance level for finite samples.

### 2.2.2 Groen and Kleibergen Test

Using iterated estimators based on the GMM approach of Hansen (1982), Groen and Kleibergen (2003) proposed LR statistics for testing cointegration in panel data. The GMM estimation and the LR test statistics depend on the

---

[30]Simulated mean and variance values of the asymptotic trace statistic can be found in Table 1 of Larsson et al. (2001).

[31]The correct proof for $d = 1$ can be found in Chapter 4, which is achieved by demonstrating the uniform boundedness of the moments of the trace statistic, i.e. $Z_{T,1}$, and the uniform integrability of $Z^2_{T,1}$.

[32]However, in the heterogeneous panel VEC model (2.135) Larsson et al. (2001) assumed that $\Pi_i = \alpha_i\beta_i'$. This means that $\Pi_i$ can differ over the cross-sections, which is conflicting with the assumptions made to establish the asymptotic distribution of the test.

following panel vector error correction model, in which Groen and Kleibergen (2003) stacked VEC models without any deterministic terms and short-run dynamics of $N$ different individuals into a joint panel VEC model.

$$\begin{aligned} \begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \vdots \\ \Delta y_{Nt} \end{pmatrix} &= \begin{pmatrix} \Pi_1 & 0 & \dots & 0 & 0 \\ 0 & \Pi_2 & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 0 & \Pi_N \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{N,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{pmatrix}, \\ \Delta Y_t &= \Pi_A Y_{t-1} + \varepsilon_t. \end{aligned} \tag{2.139}$$

$\Delta Y_t$ and $Y_{t-1}$ are of dimension $(NK \times 1)$, $\Pi_A$ is an $(NK \times NK)$ matrix and each submatrix $\Pi_i$ is $(K \times K)$ for $i = 1, \dots, N$. The $(NK \times 1)$ vector $\varepsilon_t = (\varepsilon'_{1t}, \dots, \varepsilon'_{Nt})'$ for $t = 1, \dots, T$, is a combination of $K$-dimensional error vectors of $N$ different individuals and $\varepsilon_t \sim N(0, \Omega)$, with

$$\Omega = \begin{pmatrix} \Omega_{11} & \dots & \Omega_{1N} \\ \vdots & \ddots & \vdots \\ \Omega_{N1} & \dots & \Omega_{NN} \end{pmatrix}. \tag{2.140}$$

Note that the $(K \times K)$ submatrices, $\Omega_{ij} = Cov(\varepsilon_{it}, \varepsilon_{jt}) \neq 0$ for $i, j = 1, \dots, N$. Obviously, the error terms of different cross-sections can be correlated. This assumption is an important difference from the assumption of the test of Larsson et al. (2001). While Larsson et al. (2001) estimated the VEC models with maximum-likelihood separately for each cross-section, Groen and Kleibergen (2003) used the GMM method to derive the maximum-likelihood estimators of the cointegrating vectors for the whole panel VEC model in (2.139). Additionally, with Groen and Kleibergen (2003)'s approach it is also possible to test for homogeneous long-run parameters with heterogeneous short-run dynamics.

To estimate the model with GMM-based maximum-likelihood method, $\alpha'_{i_\perp}\beta_{i_\perp}$ should have full rank for $i = 1, \dots, N$ and the true common cointegrating rank should be $r$, i.e. $\text{rank}(\Pi_i) = r$ for each $i = 1, \dots, N$, and $r < K$. Note that the estimators of $\alpha_i$, $\beta_i$ for $i = 1, \dots, N$, and $\Omega$ are fully converged estimators. The asymptotic properties of the testing procedure are investigated with the assumption that $T \to \infty$ as $N$ is fixed.

Using the common cointegrating rank assumption, the reduced form of the panel VEC model (2.139) can be rewritten as

$$\Delta Y = \begin{pmatrix} \alpha_1\beta'_1 & 0 & \dots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \dots & 0 & \alpha_N\beta'_N \end{pmatrix} Y_{t-1} + \varepsilon_t \tag{2.141}$$

$$= \Pi_B Y_{t-1} + \varepsilon_t, \tag{2.142}$$

in which $\alpha_i$ and $\beta_i$ are $(K \times r)$ matrices for $i = 1, \ldots, N$.

Applying the common cointegrating vector restriction, i.e $\beta = \beta_i$ for $i = 1, \ldots, N$, (2.142) turns into

$$\begin{aligned}\Delta Y &= \begin{pmatrix} \alpha_1\beta' & 0 & \ldots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \ldots & 0 & \alpha_N\beta' \end{pmatrix} Y_{t-1} + \varepsilon_t & (2.143)\\ &= \Pi_C Y_{t-1} + \varepsilon_t. & (2.144)\end{aligned}$$

The main interest lies on the following null and alternative hypotheses

$$H_0 : \Pi_B \quad \text{vs.} \quad H_1 : \Pi_A, \qquad (2.145)$$
$$H_0 : \Pi_C \quad \text{vs.} \quad H_1 : \Pi_A. \qquad (2.146)$$

The panel cointegration LR tests are based on the following maximum log-likelihood function, which can be defined as

$$\max \ln \ell(\widehat{\Pi}^*, \widehat{\Omega}(\widehat{\Pi^*})) = -\frac{NKT}{2}(1 + \ln(2\pi)) - \frac{T}{2} \ln |\widehat{\Omega}(\widehat{\Pi}^*)|, \qquad (2.147)$$

with $\widehat{\Pi}^* = \widehat{\Pi}_A, \widehat{\Pi}_B$ or $\widehat{\Pi}_C$. Thus, the LR test statistic for the null and alternative hypotheses (2.145) is

$$\begin{aligned}\mathrm{LR}(\Pi_B|\Pi_A) &= 2[\ln \ell(\widehat{\Pi}_A, \widehat{\Omega}(\widehat{\Pi}_A)) - \ln \ell(\widehat{\Pi}_B, \widehat{\Omega}(\widehat{\Pi}_B))] \\ &= T[\ln |\widehat{\Omega}(\widehat{\Pi}_B)| - \ln |\widehat{\Omega}(\widehat{\Pi}_A)|], & (2.148)\end{aligned}$$

and the LR test statistic for the null and alternative hypotheses (2.146) is

$$\begin{aligned}\mathrm{LR}(\Pi_C|\Pi_A) &= 2[\ln \ell(\widehat{\Pi}_A, \widehat{\Omega}(\widehat{\Pi}_A)) - \ln \ell(\widehat{\Pi}_C, \widehat{\Omega}(\widehat{\Pi}_C))] \\ &= T[\ln |\widehat{\Omega}(\widehat{\Pi}_C)| - \ln |\widehat{\Omega}(\widehat{\Pi}_A)|]. & (2.149)\end{aligned}$$

The limiting distribution of $\mathrm{LR}(\Pi_B|\Pi_A)$ as $T \to \infty$ for fixed $N$ is

$$\begin{aligned}&\mathrm{LR}(\Pi_B|\Pi_A) \xrightarrow{w} \\ &\sum_{i=1}^{N} \mathrm{tr}\left[\int_0^1 \mathrm{d}W_i(s)W_i(s)' \left(\int_0^1 W_i(s)W_i(s)'\right)^{-1} \int_0^1 W_i(s)\mathrm{d}W_i(s)'\right].\end{aligned}$$

$W_i(s)$ is a $(K - r)$-dimensional Brownian motion for individual $i$ with an identity covariance matrix. The limiting distribution of (2.149) is

$$\begin{aligned}&\mathrm{LR}(\Pi_C|\Pi_A) \xrightarrow{w} \chi^2_{((N-1)r(K-r))} \\ &+\sum_{i=1}^{N} \mathrm{tr}\left[\int_0^1 \mathrm{d}W_i(s)W_i(s)' \left(\int_0^1 W_i(s)W_i(s)'\right)^{-1} \int_0^1 W_i(s)\mathrm{d}W_i(s)'\right].\end{aligned}$$

Since the higher order dynamics only affect the short-run parameters, the limiting distributions of the statistics (2.148) and (2.149) do not change if short-run parameters (i.e. the lagged first differenced variables) are inserted into the model.

Groen and Kleibergen (2003) demonstrated that under the null hypothesis their standardized LR statistics are $N(0,1)$ distributed if $T \to \infty$ and $N \to \infty$ in such a way that $N/T \to 0$. The standardized version of the LR$(\Pi_B|\Pi_A)$ statistic is

$$\frac{[\mathrm{LR}(\Pi_B|\Pi_A)/N] - \Xi}{\sqrt{\mho}/N} \xrightarrow{w} N(0,1), \tag{2.150}$$

in which $\Xi$ and $\mho$ are the asymptotic mean and variance [33] of the limiting distribution of LR$(\Pi_B|\Pi_A)$, respectively.

The standardized LR$(\Pi_C|\Pi_A)$ for the homogeneous cointegrating vectors case can be formulated as follows.

$$\frac{[\mathrm{LR}(\Pi_C|\Pi_A)/N] - [(N-1)r(K-r)/N] - \Xi}{\sqrt{[2(N-1)r(K-r) + \mho]/N}} \xrightarrow{w} N(0,1), \tag{2.151}$$

with $\Xi$ and $\mho$ defined as above.

Finally, Groen and Kleibergen (2003) considered five different panel VEC models with a heterogenous intercept based on the following VEC model[34].

$$\Delta y_{it} = \mu_{0i} + \alpha_i \beta_i' y_{i,t-1} + \varepsilon_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \tag{2.152}$$

in which $\mu_{0i}$ is a $(K \times 1)$ vector. The five different cases Groen and Kleibergen (2003) evaluated are:

- Model 1: heterogeneous cointegrating vectors and unrestricted heterogeneous intercept, i.e. $\beta_i$ and $\mu_{0i} = c_i$ for $i = 1, \ldots, N$,

$$\Pi_{B.1} = \begin{pmatrix} \alpha_1\beta_1' & 0 & \ldots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \ldots & 0 & \alpha_N\beta_N' \\ c_1 & & \ldots & & c_N \end{pmatrix}. \tag{2.153}$$

- Model 2: heterogeneous cointegrating vectors and restricted heterogeneous intercept (the intercept lies in the cointegrating space), i.e. $\beta_i$

[33] These values are tabulated for different $K - r$ in Larsson et al. (2001).

[34] This model can also be generalized to the case with deterministic trend term.

and $\mu_{0i} = \alpha_i \delta_i'$ for $i = 1, \ldots, N$,

$$\Pi_{B.2} = \begin{pmatrix} \alpha_1\beta_1' & 0 & \ldots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \ldots & 0 & \alpha_N\beta_N' \\ \alpha_1\delta_1' & & \ldots & & \alpha_N'\delta_N \end{pmatrix}. \tag{2.154}$$

- Model 3: homogeneous cointegrating vectors and unrestricted heterogeneous intercept, i.e $\beta_i = \beta$ and $\mu_{0i} = c_i$ for $i = 1, \ldots, N$,

$$\Pi_{C.1} = \begin{pmatrix} \alpha_1\beta' & 0 & \ldots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \ldots & 0 & \alpha_N\beta' \\ c_1 & & \ldots & & c_N \end{pmatrix}. \tag{2.155}$$

- Model 4: homogeneous cointegrating vectors and restricted heterogeneous intercept, i.e $\beta_i = \beta$ and $\mu_{0i} = \alpha_i\delta_i'$ for $i = 1, \ldots, N$,

$$\Pi_{C.2} = \begin{pmatrix} \alpha_1\beta' & 0 & \ldots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \ldots & 0 & \alpha_N\beta' \\ \alpha_1\delta_1' & & \ldots & & \alpha_N\delta_N' \end{pmatrix}. \tag{2.156}$$

- Model 5: homogeneous cointegrating vectors and the restricted homogeneous intercept, i.e $\beta_i = \beta$ and $\mu_{0i} = \alpha_i\delta'$ for $i = 1, \ldots, N$.

$$\Pi_{C.3} = \begin{pmatrix} \alpha_1\beta' & 0 & \ldots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \ldots & 0 & \alpha_N\beta' \\ \alpha_1\delta' & & \ldots & & \alpha_N\delta' \end{pmatrix}. \tag{2.157}$$

The relevant LR statistics can be analogously computed. The LR tests applied to these models have the following different limiting distributions:

- Model 1:

$$\begin{aligned} &\mathrm{LR}(\Pi_{B.1}|\Pi_{A.1}) = T[\ln|\widehat{\Omega}(\widehat{\Pi}_{B.1})| - \ln|\widehat{\Omega}(\widehat{\Pi}_{A.1})|] \xrightarrow{w} \\ &\sum_{i=1}^{N} \mathrm{tr}\left[\int_0^1 \mathrm{d}W_i(s)F_i(s)' \right.\\ &\left.\left(\int_0^1 F_i(s)F_i(s)'\right)^{-1} \int_0^1 F_i(s)\mathrm{d}W_i(s)'\right] \end{aligned} \tag{2.158}$$

with

$$\Pi_{A.1} = \begin{pmatrix} \Pi_1 & 0 & \dots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \dots & 0 & \Pi_N \\ c_1 & & \dots & & c_N \end{pmatrix}, \tag{2.159}$$

and the $(K-r)$-dimensional $F_i(s)$ for each $i$ is

$$F_i(s) = \begin{pmatrix} W_i^+(s) - \int_0^1 W_i^+(s)\mathrm{d}s \\ s - \int_0^1 sds \end{pmatrix}. \tag{2.160}$$

$W_i^+(s)$ is a $(K-r-1)$-dimensional Brownian motion with identity covariance matrix and $0 \leq s \leq 1$.

- Model 2:
  The limiting distribution of

  $$\mathrm{LR}(\Pi_{B.2}|\Pi_{A.1}) \quad = \quad T[\ln|\widehat{\Omega}(\widehat{\Pi}_{B.2})| - \ln|\widehat{\Omega}(\widehat{\Pi}_{A.1})|] \tag{2.161}$$

  has the same form as in (2.158), with

  $$F_i(s) = \begin{pmatrix} W_i(s) \\ 1 \end{pmatrix}. \tag{2.162}$$

- Model 3-5: The limiting distributions of the LR statistics considering the constraints defined in (2.155)-(2.157) are

  $$\mathrm{LR}(\Pi_C^*|\Pi_{A.1}) = T[\ln|\widehat{\Omega}(\widehat{\Pi}_{C^*})| - \ln|\widehat{\Omega}(\widehat{\Pi}_{A.1})|] \xrightarrow{w} \chi^2_{(df)}$$
  $$+\sum_{i=1}^{N} \mathrm{tr}\left[\int_0^1 \mathrm{d}W_i(s)F_i(s)'\left(\int_0^1 F_i(s)F_i(s)'\right)^{-1}\int_0^1 F_i(s)\mathrm{d}W_i(s)'\right].$$

  $\widehat{\Pi}_C^* = \widehat{\Pi}_{C.1}, \widehat{\Pi}_{C.2}$ or $\widehat{\Pi}_{C.3}$ and the asymptotic distributions of the statistics are:

  - Model 3: $F_i(s)$ is the same as (2.160) and $\chi^2_{(df)} = \chi^2_{((N-1)r(K-r))}$
  - Model 4: $F_i(s)$ is the same as (2.162) and $\chi^2_{(df)} = \chi^2_{((N-1)r(K-r))}$
  - Model 5: $F_i(s)$ is the same as (2.162) and $\chi^2_{(df)} = \chi^2_{((N-1)r(K-r+1)}$

Specifications including different kinds of deterministic terms such as trends, can be derived straightforwardly. The finite sample properties of these tests are summarized in Section 2.3

### 2.2.3 Breitung Test

Breitung (2005) based his tests on a VEC model without short-run dynamics. He imposed the restriction of a homogeneous cointegrating matrix $\beta$, whereas the loading matrix $\alpha_i$ can vary over cross-sections. Breitung (2005) implemented the model suggested by Saikkonen (1999),

$$\Delta y_{it} = \alpha_i \beta' y_{i,t-1} + \phi_i \beta'_{\perp} y_{i,t-1} + \varepsilon_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T. \tag{2.163}$$

$\beta_{\perp}$ is the orthogonal complement of $\beta$ and $\phi_i$ is a $(K \times (K - r))$ matrix with full column rank. The process $\varepsilon_{it}$ is assumed to be a $K$-dimensional white noise error vector with $E(\varepsilon_{it}) = 0$ and positive definite covariance matrix $\Omega_i = E(\varepsilon_{it}\varepsilon'_{it})$. Pre-multiplying both sides of (2.163) by the orthogonal complement of $\alpha_i$, i.e. $\alpha'_{i_{\perp}}$, leads to

$$\alpha'_{i_{\perp}} \Delta y_{it} = \alpha'_{i_{\perp}} \phi_i \beta'_{\perp} y_{i,t-1} + \alpha'_{i_{\perp}} \varepsilon_{it}. \tag{2.164}$$

Let $u_{it} = \alpha'_{i_{\perp}} \Delta y_{it}$, $\delta_i = \alpha'_{i_{\perp}} \phi_i$, $w_{it} = \beta'_{\perp} y_{i,t}$ and $e_{it} = \alpha'_{i_{\perp}} \varepsilon_{it}$. The null hypothesis of no cointegration can now be formulated as

$$H_0 : \delta_i = 0 \quad \text{for all } i \quad \text{vs.} \quad H_1 : \delta_i \neq 0 \quad \text{for at least one } i. \tag{2.165}$$

For this purpose, Breitung (2005) suggested a two-step estimation procedure to estimate $\alpha_i$ and $\beta$, and thus, to obtain consistent estimates for $\alpha_{i_{\perp}}$ and $\beta_{\perp}$. This testing procedure is based on the estimation method of Ahn and Reinsel (1990) and it relies on the principle that the long-run and the short-run parameters can be estimated separately due to the fact that the Fisher information matrix is block-diagonal with respect to $\alpha_i, \Omega_i$ and $\beta$. The $\alpha_i$, $\Omega_i$ and $\beta$ matrices are estimated with the restriction that $\beta$ matrix is normalized by $\beta = (I_r, \beta_r)'$; $I_r$ is an identity matrix of dimension $r$, and $\beta_r$ is an $(r \times (K - r))$ parameter matrix. In the first step, homogeneity assumption is ignored, and the ML estimator of Johansen (1991) is used to estimate $\alpha_i$ and $\Omega_i$ for each cross-section separately. In the second step, the matrix $\beta$ is estimated by applying OLS on the following pooled regression:

$$\widehat{z}_{it} = \beta'_i y_{i,t-1} + \widehat{\nu}_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \tag{2.166}$$

with $\widehat{z}_{it} = (\widehat{\alpha}'_i \widehat{\Omega}_i^{-1} \widehat{\alpha}_i)^{-1} \widehat{\alpha}'_i \widehat{\Omega}_i^{-1} \Delta y_{it}$ and $\widehat{\nu}_{it} = (\widehat{\alpha}'_i \widehat{\Omega}_i^{-1} \widehat{\alpha}_i)^{-1} \widehat{\alpha}'_i \widehat{\Omega}_i^{-1} \varepsilon_{it}$, i.e. using the estimates for $\alpha_i$ and $\Omega_i$. Breitung (2005) also showed that the two-step estimator has a limiting normal distribution.

To test the null hypothesis $H_0 : \delta_i = 0$ for all $i$, Breitung (2005) proposed LM, LR and Wald statistics, which are denoted as $\lambda_i(r)$. The LM test statistic is derived by

$$\lambda_i^{\mathrm{LM}}(r) = T\mathrm{tr}\left[\sum_{t=1}^{T} \widehat{u}_{it}\widehat{w}'_{i,t-1}\left(\sum_{t=1}^{T} \widehat{w}_{i,t-1}\widehat{w}'_{i,t-1}\right)^{-1} \sum_{t=1}^{T} \widehat{w}_{i,t-1}\widehat{u}'_{it}\left(\sum_{t=1}^{T} \widehat{u}_{it}\widehat{u}'_{it}\right)^{-1}\right]$$

with $\widehat{u}_{it} = \widehat{\alpha}'_{i_\perp}\Delta y_{it}$ and $\widehat{w}_{it} = \widehat{\beta}'_\perp y_{it}$. The LR and Wald statistics are computed also using the consistent estimates of $\alpha_{i_\perp}$ and $\beta_\perp$. Moreover, the LR statistic can also be used to test the restrictions on the cointegrating parameters.

Breitung (2005) derived the standardized LM, LR and Wald test statistics by

$$\frac{\sqrt{N}\left[\frac{\sum_{i=1}^{N}\lambda_i(r)}{N} - E(Z_d)\right]}{\sqrt{Var(Z_d)}} \xrightarrow{w} N(0,1). \tag{2.167}$$

$Z_d$ is defined in the same way as in Larsson et al. (2001) (see Equation (2.138)). Appropriate values for $E(Z_d)$ and $Var(Z_d)$ are tabulated in Larsson et al. (2001). Note, under the null hypothesis the standardized LM, LR and Wald test statistics are asymptotically standard normal distributed if $T \to \infty$ and $N \to \infty$ sequentially.

Breitung (2005) assumed that the test statistics can be generalized to a cointegrated VAR($p$) process with deterministic terms. Then, the underlying asymptotic distribution of the test statistics will depend on the deterministic terms included in the model. Thus, for three different specifications of the deterministic terms[35], Breitung (2005) tabulated the corresponding asymptotic mean and variance values to standardize the test statistics. He assumed that the standardized versions of these tests have standard normal distributions as $T$ and $N$ go to infinity sequentially. However, he did not provide a proof on the finiteness of the asymptotic moments, which is necessary to establish the asymptotic distributions of these tests.

In a simulation study Breitung (2005) demonstrated that his tests have better size and power properties than the test of Larsson et al. (2001). For small samples, the sizes of Breitung's tests are around the 5% significance level.

[35] Model 1: the variables have no time trend, Model 2: at least one variable has a linear time trend, Model 3: the variables and the cointegrating relations have a linear time trend.

### 2.2.4 Anderson, Qian and Rasche Test

Anderson et al. (2006) proposed to use the canonical correlation method of Box and Tiao (1977) to test for cointegrating rank in panel VEC models. Anderson et al. (2006) demonstrated that, under cross-sectional correlation and cross-sectional cointegration the panel cointegration tests, which do not allow for these characteristics, will suffer from severe size distortions and low power. They proposed a test that accommodates for cross-sectional dependence between the shocks, cross-sectional dependencies in short-run dynamics, differences in cointegrating rank in cross-sections, and finally the existence of long-run equilibrium relationships between different cross-sections.

They considered the panel VEC model of the $K$-dimensional vector $y_{it}$:

$$\begin{aligned}\Delta y_{it} &= C_i d_t + \Pi_i y_{i,t-1} + \sum_{j=1}^{p-1} \Gamma_{ij} \Delta y_{i,t-j} + \varepsilon_{it}, \\ & \quad i = 1, \ldots, N; \quad t = 1, \ldots, T.\end{aligned} \tag{2.168}$$

$\Pi_i$ and $\Gamma_{ij}$ are defined in the same way as in (2.135). However, the order of the VEC process do not vary over cross-sections, i.e. $p_i = p$. Moreover, $d_t$ is either $d_t = 1$ or $d_t = (1, t)'$ and $C_i$ is the relevant unknown parameter vector or matrix depending on the dimension of $d_t$.

Anderson et al. (2006) rewrote the model in (2.168) as

$$\Delta y_{it} = C_i d_t + \Pi_i y_{i,t-1} + \Gamma_i X_{it} + \varepsilon_{it}, \tag{2.169}$$

with

$$\Gamma_i = (\Gamma_{i1}, \Gamma_{i2}, \ldots, \Gamma_{i,p-1}), \tag{2.170}$$
$$X_{it} = (\Delta y'_{i,t-1}, \Delta y'_{i,t-2}, \ldots, \Delta y'_{i,t-(p-1)})'. \tag{2.171}$$

Using the reduced rank decomposition of $\Pi_i$, they proposed the following model.

$$
\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \vdots \\ \Delta y_{Nt} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} d_t + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1N} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{N1} & \alpha_{N2} & \dots & \alpha_{NN} \end{bmatrix}
\begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1N} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{N1} & \beta_{N2} & \dots & \beta_{NN} \end{bmatrix}' \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{N,t-1} \end{bmatrix} +
\begin{bmatrix} \Gamma_{11} & \Gamma_{12} & \dots & \Gamma_{1N} \\ \Gamma_{21} & \Gamma_{22} & \dots & \Gamma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{N1} & \Gamma_{N2} & \dots & \Gamma_{NN} \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{Nt} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix} \tag{2.172}
$$

In compact notation

$$
\Delta y_t = C d_t + \alpha\beta' y_{t-1} + \Gamma X_t + \varepsilon_t, \quad \text{for} \quad t = 1, \dots, T. \tag{2.173}
$$

$\alpha$ and $\beta$ are $(NK \times r)$ matrices with $r = r_1 + r_2 + \ldots + r_N < NK$ and $y_t = (y'_{1t}, y'_{2t}, \ldots, y'_{Nt})'$. Representation (2.172) allows for interaction of short-run dynamics between cross-sections, different cointegrating rank across cross-sections and cross-sectional cointegration. Anderson et al. (2006) assumed that the *i.i.d.* equilibrium error process $\varepsilon_t$ has a zero mean vector and the following covariance matrix.

$$
\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} & \dots & \Omega_{1N} \\ \Omega_{21} & \Omega_{22} & \dots & \Omega_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Omega_{N1} & \Omega_{N2} & \dots & \Omega_{NN} \end{bmatrix}, \tag{2.174}
$$

which is an $(NK \times NK)$ positive definite matrix, with $\Omega_{ij} = E(\varepsilon_{it}\varepsilon'_{jt})$ for $i, j = 1, \ldots, N$. In contrast to the form in Larsson et al. (2001), this form of the covariance matrix accommodates for cross-sectional dependence.

Anderson et al. (2006) implemented bootstrap method of Giersbergen (1996) to find the empirically correct finite sample distributions of the cointegrating rank tests. This is done, because this method does not require any distributional assumptions about the DGP. The method of Box and Tiao (1977) is built on the predictability of linear combinations of multivariate time series from the history of the linear combinations. A stationary linear combination has no predictive power for the current value. On the contrary, if the linear combination of the multivariate time series is nonstationary, then its past values help to forecast the current value. Suppose $y_t = \widehat{y}_t + \widehat{e}_t$ and $\widehat{y}_t = \Gamma_1 y_{t-1} + \ldots + \Gamma_p y_{t-p}$ is the linear projection of $y_t$ on its own history with the $(K \times K)$ projection coefficient matrices $\Gamma_j$'s. $\widehat{e}_t$ is the projection error, which is uncorrelated with $\widehat{y}_t$. Anderson et al. (2006) defined

$$\begin{aligned}
\Lambda &= \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_K), \quad \text{with} \quad 0 \leq \lambda_i \leq 1 \quad \text{for} \quad i = 1, \ldots, K, \\
V &= (\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_K)', \\
z_t &= \begin{bmatrix} \mathrm{v}_1' y_t \\ \mathrm{v}_2' y_t \\ \vdots \\ \mathrm{v}_K' y_t \end{bmatrix} = V y_t, \\
z_t &= \widehat{z}_t + \widehat{q}_t.
\end{aligned} \tag{2.175}$$

$\Lambda$ is the diagonal matrix of eigenvalues of $Cov(\widehat{y}_t)$ in the metric of $Cov(y_t)$, $V$ is the matrix of eigenvectors, $z_t$ represents the canonical vector, $\widehat{z}_t = V\widehat{y}_t$ and $\widehat{q}_t = V\widehat{e}_t$. (2.175) implies that

$$Cov(z_t) = Cov(\widehat{z}_t) + Cov(\widehat{q}_t) \tag{2.176}$$

because $\widehat{z}_t$ and $\widehat{q}_t$ are independent. After the normalization with $\mathrm{v}_i' Cov(y_t)\mathrm{v}_i = 1$, for $i = 1, \ldots, K$ it follows that $Cov(z_t) = I_K$, $Cov(z_t, \widehat{z}_t) = Cov(\widehat{z}_t) = \Lambda$ and $Cov(\widehat{q}_t) = I_K - \Lambda$. Box and Tiao (1977) showed that the eigenvalues[36] $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_K$ are just the squared canonical correlations between $y_t$ and $\widehat{y}_t$. In other words, if the largest eigenvalue $\lambda_K$ is close to 1, then the canonical covariate $z_{Kt} = \mathrm{v}_K' y_t$ is $I(1)$ and is highly predictable. In contrast, if smallest eigenvalue $\lambda_1$ is significantly less than one, then $z_{1t} = \mathrm{v}_1' y_t$ is $I(0)$ and not predictable. Briefly, the number of eigenvalues close to 1 gives us the number of stochastic common trends, i.e. $K - r$. As a result, the null and alternative hypotheses to test for the number of common stochastic trends can be formulated as

$$H_0 : \lambda_{r+1} = 1 \quad \text{vs.} \quad H_1 : \lambda_{r+1} < 1, \tag{2.177}$$

with $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_K$. To test the null hypothesis a DF or an ADF type unit root test on $z_{it}$ can be implemented.

Additionally, in an empirical example on M1 velocities, Anderson et al. (2006) suggested to apply bootstrapping techniques to the Johansen trace statistic[37] because the asymptotic distribution of Johansen's trace statistic

[36]Please note that $\lambda_i = \frac{\mathrm{v}_i' Cov(\widehat{y}_t)\mathrm{v}_i}{\mathrm{v}_i' Cov(y_t)\mathrm{v}_i} = \frac{Cov(\widehat{\mathrm{v}_i' y_t})}{Cov(\mathrm{v}_i' y_t)} = \frac{Cov(\widehat{z}_{it})}{Cov(z_{it})}$.

[37]Johansen trace statistic is based on the canonical correlation of Bewley and Yang (1995).

is not a good approximation to the true distribution of the test statistic for small and moderate sample sizes[38].

### 2.2.5 Larsson and Lyhagen Test

Larsson and Lyhagen (1999) suggested two LR panel cointegration tests to test for the cointegrating restrictions in the panel VEC models. Their tests are based on a similar model like (2.172) without deterministic terms, but with an additional restriction that the $\beta$ matrix is block diagonal. Thus, similar to (2.173), the model can be compactly written as

$$\Delta y_t = \alpha\beta' y_{t-1} + \Gamma X_t + \varepsilon_t, \tag{2.178}$$

with

$$\beta = \begin{bmatrix} \beta_{11} & 0 & \dots & 0 \\ 0 & \beta_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \beta_{NN} \end{bmatrix}. \tag{2.179}$$

The $NK$-dimensional vector $\varepsilon_t \sim N(0, \Omega)$ and $\Omega$ has the same structure as (2.174) and the $\Gamma$ matrix is defined as in (2.172). (2.178) and (2.174) represent a generalized version of the model defined in Larsson et al. (2001). Since $\alpha$ is not block-diagonal, the model allows for short-run dependencies between the cross-sections. The cointegrating relations are only allowed within the cross-section because the $\beta$ matrix is restricted to be block-diagonal. There is also cross-sectional correlation and interaction between the cross-sections through the short-run coefficients as the off-diagonal elements of $\Omega$ and $\Gamma$ are not zero. Moreover, each cross-section may have a different number of cointegrating relations.

Larsson and Lyhagen (1999) considered two different homogeneity hypotheses. With the following first pair of hypotheses

$$H_0 : \text{rank}(\Pi) = r_i \leq r \text{ vs. } H_1 : \text{rank}(\Pi) = K \text{ for } i = 1, \dots, N, \tag{2.180}$$

they wanted to test that all of the cross-sections have at most $r$ cointegrating relations. The unknown parameters of the restricted and unrestricted models are estimated following Johansen (1995). If $r > 0$ and $\alpha'_\perp \Gamma \beta_\perp$ has full rank, the variables are not integrated of a higher order than one. The LR statistic for the hypothesis (2.180) has a limiting distribution, which is a combination of an $N(K - r)$-dimensional standard Brownian motion functional with an

[38]see Toda, 1995; Haug, 1996 for their simulation studies.

identity covariance matrix and a $\chi^2$ distribution with $N(N-1)(K-r)r$ degrees of freedom.

After the determination of the cointegrating rank, Larsson and Lyhagen (1999) proposed a test for a second pair of hypotheses to test whether the cointegrating vectors $\beta_{ii}$ span the same space.

$$H_0 : \beta_{11} = \beta_{22} = \ldots = \beta_{NN} = \beta \text{ vs. } H_1 : \beta_{ii} \neq \beta_{jj} \text{ for some } i, j. \quad (2.181)$$

Hence, under the null hypothesis $\beta = (I_N \otimes \beta)$. In other words, with the hypotheses in (2.181), the restricted model $\beta = (I_N \otimes \beta)$ is tested against the model $\beta = \mathrm{diag}(\beta_{ii})$ using an LR statistic. Consequently, under the null hypothesis LR statistic is asymptotically $\chi^2$ distributed with $(N-1)r(K-r)$ degrees of freedom. To estimate the unknown coefficients of the restricted model, Larsson and Lyhagen (1999) implemented the switching algorithm of Boswijk (1996).

The test for the cointegrating rank determination reveals size distortions in finite samples. To correct for the bad size properties, Larsson and Lyhagen (1999) suggested using Bartlett corrected test statistic,

$$C_T^* = E(C_T)\frac{C_\infty}{E(C_\infty)},$$

in which $C_T$ is the test statistic for a sample size $T$, $C_\infty$ represents the asymptotic test statistic and $E$ is the expectation operator.

Please note that estimation of the parameters is infeasible when the cross-section dimension increases. Then, the system is not identifiable anymore. Moreover, the asymptotic distributions of the statistics are derived under the assumption that $T \to \infty$ as $N$ is held fixed and small.

## 2.3 Finite Sample Properties

In this section a review of the Monte Carlo studies performed by several authors to compare the finite sample properties of the panel cointegration tests will be given. Most of the simulation studies demonstrates that the usage of panel data cointegration techniques increases the power in comparison to the conventional cointegration techniques. Generally, the tests have high power if the assumptions of the tests are fulfilled. Moreover, for small time dimensions the tests exhibit size distortions, but the sizes of the tests approaches the nominal size level with the increase in $T$.

The first known extensive simulation study was performed by McCoskey and Kao (1999). They examined the size and size-adjusted power properties of the four residual-based null hypothesis of no cointegration tests[39] and the panel LM null hypothesis of cointegration test of McCoskey and Kao (1998). The main finding of their study is that the size of the tests can be affected by the relative sizes of $T$ and $N$, however, ADF* has the best size properties. Generally, the power increases more with an increase in $T$ than in $N$. According to the results of McCoskey and Kao (1999), the panel LM test outperforms all the other tests and is the most powerful one, even in the presence of a nearly nonstationary data[40]. As a result, they concluded that it is more appropriate to test for the null hypothesis of cointegration instead of the null hypothesis of no cointegration.

Gutierrez (2003) compared the powers of the five tests of Kao (1999), the seven tests of Pedroni (1999) and the maximum-likelihood-based test of Larsson et al. (2001)[41]. Gutierrez (2003) used finite sample critical values, instead of asymptotic ones and he modified the DGP applied by Engle and Granger (1987), Gonzalo (1994) and Haug (1996) to panel data. Gutierrez (2003) took into account that not the whole panel, but only a fraction of the panel is cointegrated and the vector of regressors is weakly exogenous. The main outcome is that the tests of Kao (1999) and Pedroni (1999) perform better than the standardized LR-bar statistic of Larsson et al. (2001). Kao's $\mathrm{DF}^*_\rho$ and Pedroni's group-$\rho$ and panel-$\rho$ tests have the highest power. On the one hand, if the time dimension of the panel is small, the tests of Kao (1999) outperform the tests of Pedroni (1999). On the other hand, if the time dimension is large the tests of Pedroni have higher power than those of Kao. The powers of the tests increase when $N$ increases for fixed $T$, but the increase in power is higher if $T$ increases while $N$ is fixed. Gutierrez (2003) also pointed out the fact that for panels with small $T$ there is a risk in concluding that the panels are not cointegrated although a high fraction of the panel is cointegrated. On the contrary, for panels with large $T$ there is a risk to conclude that the panels are cointegrated although a small fraction of the panel is cointegrated. The results do not change when the vector of regressors is endogenous.

Groen and Kleibergen (2003) compared their test statistic $\mathrm{LR}(\Pi_{B.2}|\Pi_{A.1})$, which accommodates for cross-sectional dependence, with the standardized LR-bar test of Larsson et al. (2001). They showed that the latter test is severely size distorted if it is implemented to a panel data with cross-sectional dependence. In addition to this, the tests show gain in power, when they are

[39]The standardized average Augmented Dickey-Fuller test (ADF*), the standardized average Phillips $\mathrm{Z}_t$ test ($\mathrm{PO}^*_t$), the standardized panel version of Phillips and Ouliaris-$\rho$ test ($\mathrm{PO}^*_\alpha$) and the standardized panel-$t$ test of Pedroni (1999) (APG*).

[40]The parameter of the AR(1) process representing the residuals is near unity.

[41]Please note that all these tests rely on the independence of the cross-sections.

considered in a panel cointegration framework than the usual conventional time series framework.

Banerjee et al. (2004) performed an extensive simulation study to figure out the finite sample properties of the tests of Larsson and Lyhagen (1999) and Pedroni (1999, 2004) given cross-sectional cointegration within the panel. Briefly, cross-sectional cointegration defines the case when the $\beta$ matrix in (2.179) is not block-diagonal anymore. The simulation results state that the Larsson and Lyhagen (1999) test interprets the presence of cross-sectional cointegration as cointegration within the cross-section, and overrejects the null hypothesis. However, the presence of cross-sectional cointegration is also a problem for large $N$. Furthermore, if there is no cross-sectional cointegration, the test of Larsson and Lyhagen (1999) has reasonable size and power, but it becomes size distorted with an increase in $N$. Banerjee et al. (2004) also revealed that the Bartlett-correction removes size distortions and increases the power of the Larsson and Lyhagen (1999) test. In the presence of cross-sectional cointegration, the tests of Pedroni (1999) overrejects the null hypothesis of no cointegration, mainly for small $T$ and large $N$. However, the distortions are lower than the Larsson and Lyhagen (1999) test if $N$ is small. Without cross-sectional cointegration the tests of Pedroni also exhibit size distortions, when $N$ increases for small $T$. Moreover, generally the size distortions decrease with the increase in $T$ for fixed $N$. Before testing for panel cointegration, Banerjee et al. (2004) recommended to test first, whether the cointegrating rank is homogeneous[42] over cross-sections, and then to test for no cross-sectional cointegration using Gonzalo and Granger (1995) test.

Hanck (2007) compared the power properties of his tests i.e. $P_{\chi^2}$, $P_{\Phi^{-1}}$, $P_t$, with the results from the simulation study of Gutierrez (2003) outlined above. Both authors implemented the DGP of Engle and Granger (1987) on panel data with the same parameters. Unfortunately, Hanck (2007) did not consider size-adjusted power. He observed that for shorter panels the tests of Pedroni (2004) and Kao (1999) have higher power than the tests of Hanck (2007). The simulation results point out that the $p$-value based tests of Hanck always outperform the standardized LR-bar test of Larsson et al. (2001). With a second DGP, Hanck introduced heterogeneity in the serial correlation of the error terms by modelling the equilibrium errors of the system as an AR process, in which the lag order is allowed to vary over the cross-sections. However, simulation study with this DGP shows that Kao's

[42]Larsson and Lyhagen (1999) and Pedroni (1999, 2004) assumed that the cointegrating rank is homogeneous over cross-sections. Banerjee et al. (2004) demonstrated also that these tests suffer from size distortions if this assumption is violated.

tests have sincerely less power than the tests of Pedroni and Hanck that do not accommodate for heterogeneity.

In another study, Hanck (2006a) analyzed the consistency in the decision of the various tests proposed in panel cointegration literature. He reexamined the tests introduced in Hanck (2007), Pedroni (2004), Kao (1999) and Larsson et al. (2001). Hanck (2006a)[43] implemented the method designed by Gregory et al. (2004), i.e calculated the correlation of the $p$-values for pairs of statistics, and observed the pairs of tests which reject together. The striking outcome is that different pairs of panel cointegration tests give conflicting results for the same dataset. There is even no high correlation between the pairs of tests proposed by the same author.

This outcome is also similar to the results of Westerlund and Basher (2008). They investigated a different problem of the panel cointegration techniques. Most of the residual-based panel cointegration tests, either parametric or semi-parametric, deal with the problems of bandwidth and kernel estimator selection and the determination of the appropriate lag order of the AR process. Westerlund and Basher (2008) examined the tests of Pedroni (1999, 2004) and Westerlund (2005a), to analyze the consistency of the decision of the tests with respect to different adjustment methods for the temporal dependencies. As it was explained earlier, the tests of Pedroni (1999, 2004) require the appropriate bandwidth and lag order selection. On the contrary, the panel cointegration tests of Westerlund (2005a) are nonparametric and do not require any lag order or bandwidth selection. Westerlund and Basher (2008) considered two different modelling of the equilibrium error terms[44]. The DGP allows for endogenous regressors and deterministic terms. For comparison different truncation windows and lag order selection methods are applied. They demonstrated that different adjustment methods do not reveal a clear result for the tests of Pedroni (1999, 2004). The sizes of the tests of Westerlund (2005a) are at around the 5% significance level for longer panels, but they are also as size distorted as the tests of Pedroni when the AR(1) or the MA(1) coefficients of the equilibrium error terms are negative[45]. Moreover, the tests of Pedroni (1999, 2004) suffer from size distortions even when the AR(1) and MA(1) coefficients are zero.

Westerlund (2005a) also compared the tests of Pedroni (1999, 2004) with his tests proposed in Westerlund (2005a). In contrast to Westerlund and Basher (2008), he considered the case in which the error terms are driven

[43]He employed again the DGP used by Gutierrez (2003), and calculated the $p$-values for the ADF test of Engle/Granger and the trace test of Johansen.

[44]In the first case the error terms follow an AR(1) process. In the second case they follow an MA(1) process.

[45]Sometimes the sizes of the tests of Pedroni (1999, 2004) approach unity.

by common factors to obtain a panel with cross-sectional dependence. In this scenario the tests of Pedroni always reject the true null hypothesis. The variance ratio tests of Westerlund have better size properties with small distortions when the DGP has a common time effects representation[46]. Westerlund's tests are also size distorted if the loading factors are $\lambda_i \sim N(1,1)$, but the distortion is not as severe as for the tests of Pedroni[47]. For $\mathrm{VR}_P$ and $\mathrm{VR}_G$, the size distortions decrease if $T$ increases. The size-adjusted power results reveal that the panel mean tests are more powerful than the group mean tests. Among the panel mean tests, the $\mathrm{VR}_P$ test of Westerlund (2005a) is the second most powerful test after Pedroni's variance ratio test. Moreover, $\mathrm{VR}_G$ is the most powerful test among the group mean tests. As expected, the powers of all the tests approach unity with the increase in $T$ and $N$.

Westerlund (2005b) conducted another Monte Carlo study to compare his panel CUSUM test for the null hypothesis of cointegration with the panel LM null hypothesis of cointegration test of McCoskey and Kao (1998). The study is built on the design of Xiao and Phillips (2002). The DGP allows for serial correlation, deterministic terms and endogenous regressors. Westerlund (2005b) also used two different estimation methods: FMOLS and DOLS. The simulation results show that, in general, the panel CUSUM test has better size than the panel LM test and that there is not much difference between the results based on the DOLS and FMOLS methods. Both tests are oversized when the autoregressive parameter $\rho$, which shows the persistence in the DGP, is $\rho \geq 1.5$. However, the distortion in the panel LM test is more severe. On the contrary, when there is no persistence in the DGP, i.e $\rho = 0$, the panel CUSUM test is undersized, whereas the empirical size of the panel LM test is around the 5% level. If the fraction of the integrated equilibrium errors in the panel is low, both tests have low size-adjusted power. The power increases under two conditions; if $T$ and $N$ grow, and if the fraction of the integrated equilibrium errors is high.

Gengenbach et al. (2006) demonstrated that the popular panel cointegration test statistics of Kao (1999) and Pedroni (2004) do not have a limiting normal distribution. The tests suffer from size distortions if the underlying assumption of cross-section independence is violated.

Westerlund (2007) examined the finite sample properties of his error correction based cointegration tests and the residual-based tests of Pedroni (1999, 2004). First assuming that there is no cross-sectional dependence,

[46]The case, when the loading factor $\lambda_i = 1$. Then, the size is at most 7.4%, for $T = 50$, $N = 20$ without intercept in the DGP and at least 2.8% for $T = 50$, $N = 20$ with a linear time trend in the DGP.

[47]The sizes of Westerlund's tests are at most 48.6%, for $T = 50$, $N = 20$ with a linear time trend in the DGP.

Westerlund (2007) investigated the size properties of the tests, given that the equilibrium errors follow an MA(1) process, and there is weak exogeneity. In the presence of serial correlation or when the strict exogeneity of the regressors is violated, the tests of Pedroni are severely size distorted. Instead, the error correction tests show relatively less size distortions. Among the tests of Westerlund (2007), $\mathrm{G}_\tau$ and $\mathrm{P}_\tau$ perform better than $\mathrm{G}_\alpha$ and $\mathrm{P}_\alpha$. Error correction tests have higher power than the residual-based tests of Pedroni. Hence, the power increases with $T$, $N$ and the signal-to-noise ratio. Under cross-sectional dependence, the bootstrap versions of $\mathrm{G}_\tau$, $\mathrm{P}_\tau$, $\mathrm{G}_\alpha$ and $\mathrm{P}_\alpha$ have better size properties than the tests of Pedroni, while the tests of Pedroni are severely size distorted. The bootstrap tests have higher power in comparison to the tests of Pedroni, and the panel bootstrap statistics are more powerful than the group bootstrap statistics. Moreover, Westerlund (2007) did not suggest using error correction panel cointegration tests if weak exogeneity of regressors is violated, as they may not work properly for some cases.

Wagner and Hlouskova (2007) considered the panel cointegration tests proposed by Pedroni (1999, 2004), Westerlund (2005a), Larsson et al. (2001) and Breitung (2005). To observe the finite sample properties of these panel cointegration tests their DGP allows for autoregressive roots near unity, short-run cross-sectional dependence, cross-sectional cointegration and $I(2)$ components. The DGP is based on a three-dimensional VAR(2) process. Three different methods are considered to formulate the short-run cross-sectional correlation: constant correlation[48], the Toeplitz correlation[49] and a common factor modelling that consists of two stationary factors. According to the results, the residual-based tests of Pedroni (panel-$t$ and group-$t$ tests) have the best performance. They are less affected by the presence of an $I(2)$ component, cross-sectional dependence or cross-sectional cointegration, whereas the tests of Westerlund (2005a) are undersized. Unfortunately, the maximum-likelihood-based tests are sensitive to the presence of $I(2)$ components which has a strong negative impact. Additionally, the systems cointegration tests have a tendency to choose a higher rank than the true rank. They have poor properties for small $T$, even when $N$ increases. On the contrary, the systems cointegration tests are not affected by the presence of autoregressive roots near unity as the residual-based tests are. Wagner and Hlouskova (2007) suggested using the ADF type tests of Pedroni (1999, 2004) if the aim is to test the null hypothesis of no cointegration because

[48] The correlation between any two series from different cross-sections is $\kappa$ times larger than the correlation between the same kind of two series from the same cross-section; $\kappa$ is some constant.

[49] In this method the correlations decreases geometrically with distance.

the panel-$t$ and group-$t$ statistics outperform the maximum-likelihood-based tests. To correct for the poor properties of the maximum-likelihood-based tests and the tests of Westerlund (2005a), Wagner and Hlouskova (2007) recommended to use finite sample critical values or bootstrap methods.

Gutierrez (2008) compared his tests, which accommodate for structural shifts, with the non-parametric panel-$t$ and group-$t$ statistics of Pedroni (1999, 2004). The simulation results reveal that the tests of Pedroni have almost no power if the presence of a break in the time trend or in the cointegrating vector is ignored. On the contrary, the presence of a level shift is not so severe if the break is at the beginning of the sample, and $T$ and $N$ are large. If there is no break, especially the size of the group-$t$ test is around the nominal 5% level and both tests of Pedroni are highly powerful compared to the tests of Gutierrez (2008). However, in the presence of structural breaks the latter tests are powerful for moderate and large panels and are not influenced by the location of the break.

# Chapter 3

# Comparison of Panel Cointegration Tests

## 3.1 Introduction

In this chapter two within-dimension-based (panel-$\rho$ and parametric panel-$t$) and two between-dimension-based (group-$\rho$ and parametric group-$t$) panel cointegration tests of Pedroni (1999) are compared with the maximum-likelihood-based panel cointegration test of Larsson et al. (2001). The group-$\rho$ test is chosen because Gutierrez (2003) demonstrated that this test has the best power among the tests of Pedroni (1999), Larsson et al. (2001) and Kao (1999). The parametric group-$t$ test is selected as the data generating process, which I use for the simulation study, is appropriate for parametric ADF type tests. Additionally, the within-dimension versions of these tests (i.e. panel-$\rho$ and parametric panel-$t$) are taken into account to be able to compare them with their between-dimension versions.

For the tests of Pedroni (1999) the regression equation with a heterogeneous intercept is considered. Note that as introduced in Chapter 2 with Equation (2.1), the regression equation could also be estimated without a heterogeneous intercept, or with a time trend and/or common time dummies. To determine the lag truncation order of the ADF $t$-statistics, the step-down procedure and the Schwarz lag order selection criterion are used.

As I explained in Chapter 2, Larsson et al. (2001) presented a maximum-likelihood-based panel test for the cointegrating rank in heterogeneous panels. They proposed a standardized LR-bar test based on the mean of the individual rank trace statistic of Johansen (1995).

The outline of this chapter is as follows: Section 3.2 introduces the DGP of Toda (1994, 1995) which is modified for panel data, and Section 3.3 gives a description of the simulation study. Section 3.4 provides the simulation results. The validity of the Fisher relation is the subject of the Section 3.5. Conclusions are given in Section 3.6.

## 3.2 Data Generating Process

The Monte Carlo study is based on the data generating process of Toda (1994, 1995), which has been used in several studies[1]. The canonical form of the Toda process allows to see the dependence of the test performance on some key parameters.

Let $y_{it}$ be a $K$-dimensional vector, where $i$ is the index for the cross-section, $t$ is the index for the time dimension and $K$ denotes the number of variables in the model. The data generating process has the form of a VAR(1) process. The general form of the modified Toda process for a system of three variables in the absence of a linear time trend in the data is

$$y_{it} = \begin{pmatrix} \psi_a & 0 & 0 \\ 0 & \psi_b & 0 \\ 0 & 0 & \psi_c \end{pmatrix} y_{i,t-1} + \varepsilon_{it}, \qquad i = 1, \ldots, N; \qquad t = 1, \ldots, T. \tag{3.1}$$

The initial values of $y_{it}$ vector, which can be represented as $y_{i0}$ are zero. The error terms for each cross-section have the following structure.

$$\varepsilon_{it} = \begin{pmatrix} \varepsilon_{1it} \\ \varepsilon_{2it} \end{pmatrix} \sim N\left(0, \begin{pmatrix} I_r & \mathbf{\Theta} \\ \mathbf{\Theta}' & I_{K-r} \end{pmatrix}\right) \; i.i.d. \tag{3.2}$$

The cointegrating rank of the process is denoted by $r$ and $\varepsilon_{1it}$, $\varepsilon_{2it}$ are the disturbances to the stationary and nonstationary components of the DGP, respectively. $\mathbf{\Theta}$ represents the vector of instantaneous correlations between the stationary and nonstationary components of the relevant cross-section.

Taking into account (3.1), if $\psi_a = \psi_b = \psi_c = 1$, a cointegrating rank of zero is obtained. Thus, the DGP becomes

$$y_{it} = I_3 y_{i,t-1} + \varepsilon_{it} \quad \text{with} \quad \varepsilon_{it} \sim N(0, I_3) \; i.i.d. \tag{3.3}$$

This shows that the process consists of three nonstationary components and these components are instantaneously uncorrelated. The VEC representation of (3.3) is

$$\Delta y_{it} = \Pi_i y_{i,t-1} + \varepsilon_{it}. \tag{3.4}$$

[1] Lütkepohl and Saikkonen (2000), Saikkonen and Lütkepohl (1999, 2000b), Hubrich et al. (2001), Trenkler (2008), Trenkler et al. (2008) etc.

In (3.4), $\Pi_i = -(I_3 - A_{i1})$, and $A_{i1} = I_3$ represents the coefficient matrix of the VAR(1) process from (3.3). Since $\Pi_i$ is a null matrix, (3.4) turns into

$$\Delta y_{it} = \varepsilon_{it}.$$

With $|\psi_a| < 1$ and $\psi_b = \psi_c = 1$ the true cointegrating rank of the DGP is one, and it is composed of one stationary and two nonstationary components. The DGP can be formulated as

$$y_{it} = \begin{pmatrix} \psi_a & 0 \\ 0 & I_2 \end{pmatrix} y_{i,t-1} + \varepsilon_{it} \quad \text{with} \quad \varepsilon_{it} \sim N\left(0, \begin{pmatrix} 1 & \mathbf{\Theta} \\ \mathbf{\Theta}' & I_2 \end{pmatrix}\right) \; i.i.d., \tag{3.5}$$

$\mathbf{\Theta} = (\theta_a, \theta_b)$ and $|\theta_a|, |\theta_b| < 1$.

The cointegrating rank of the process (3.1) is two if $\psi_a$ and $\psi_b$ are less than unity in absolute values and $\psi_c = 1$. This can be represented in matrix form by

$$y_{it} = \begin{pmatrix} \psi_a & 0 & 0 \\ 0 & \psi_b & 0 \\ 0 & 0 & 1 \end{pmatrix} y_{i,t-1} + \varepsilon_{it} \quad \text{with} \quad \varepsilon_{it} \sim N\left(0, \begin{pmatrix} I_2 & \mathbf{\Theta}' \\ \mathbf{\Theta} & 1 \end{pmatrix}\right) \; i.i.d., \tag{3.6}$$

$\mathbf{\Theta} = (\theta_a, \theta_b)$ and $|\theta_a|, |\theta_b| < 1$. The process (3.6) consists of two stationary and one nonstationary component. These components are correlated if at least $\theta_a$ or $\theta_b$ is different from zero.

If $|\psi_a|, |\psi_b|$ and $|\psi_c| < 1$, the DGP is an $I(0)$ process with the true cointegrating rank three, which can be represented as

$$y_{it} = \begin{pmatrix} \psi_a & 0 & 0 \\ 0 & \psi_b & 0 \\ 0 & 0 & \psi_c \end{pmatrix} y_{i,t-1} + \varepsilon_{it} \quad \text{with} \quad \varepsilon_{it} \sim N(0, I_3) \; i.i.d. \tag{3.7}$$

## 3.3 Simulation Study

In order to the see how the performance of the tests is affected by some key parameters, throughout the simulation study the time and cross-section dimensions, the parameters $\psi_a, \psi_b, \psi_c$ and the correlation parameters $\theta_a$ and $\theta_b$ vary.

The correlation parameters $\theta_a$ and $\theta_b$ take on the values $\{0, 0.4, 0.7\}$. The $\psi$ parameters take on the values $\{0.5, 0.8, 0.95, 1\}$. $\psi = 0.95$ is chosen to see how the tests react when the true cointegrating rank of the process is near zero. The performance of the tests under the assumption of no instantaneous correlation between the disturbances is checked by $\theta_a = \theta_b = 0$.

To compare the results with Larsson et al. (2001), for the cross-section dimension I employ $N \in \{1, 5, 10, 25, 50\}$ and for the time dimension $T \in \{10, 25, 50, 100, 200\}$, using 1000 number of replications. While generating the random error terms, seeded values are used and the first 100 observations are deleted.

The maximum lag order for the panel-$t$ and group-$t$ statistics is limited to 3 because this is the maximum lag order allowing an efficient estimation for small time dimensions, e.g $T = 10$. To select the lag order for the panel-$\rho$ and group-$\rho$ statistics a kernel estimator is used as explained in Chapter 2. For the maximum-likelihood-based test statistic no VAR model lag order selection criterion is considered because the data is generated using a VAR(1) process. Only the null hypothesis of no cointegration hypothesis is tested as the residual-based tests cannot test for the order of panel cointegrating rank. All the results presented below are obtained by GAUSS 5.0.

## 3.4 Simulation Results

The most interesting simulation results are presented in Tables[2] 3.1 to 3.7. The size and size-adjusted power properties of the tests are considered separately in the next two subsections.

### 3.4.1 Empirical Size Properties

Table 3.1 summarizes the results for the empirical size of the tests. It is obvious that the empirical sizes of the group-$\rho$ and panel-$\rho$ tests are almost always zero for $T = 10$, 25 and $N \geq 1$, which means that the true hypothesis of no cointegration can never be rejected. The severe size distortions for the other tests when $T$ is small and $N$ is large, can also be easily recognized (e.g. the empirical sizes of the tests, except for the group-$\rho$ and panel-$\rho$ tests, are unity if $T = 10$ and $N \geq 25$). Moreover, it can be concluded that, for $T < 200$ the tests become more oversized with increasing $N$. These tests are not appropriate if the time dimension is much smaller than the cross-section dimension. The reason for this may be the fact that I use the asymptotic first and second moments in order to standardize the test statistics. Thus, the appropriate moments from the finite sample distribution of the test statistics should be used for

[2] In the tables, “sc" is the abbreviation for the Schwarz lag selection criterion and “sd" denotes the step-down lag selection method.

**Table 3.1:** Empirical size results of the tests.

| T | N | group-$\rho$ | panel-$\rho$ | group-$t$(sd) | group-$t$(sc) | panel-$t$(sd) | panel-$t$(sc) | std. LR-bar |
|---|---|---|---|---|---|---|---|---|
| 10 | 1 | 0.000 | 0.000 | 0.349 | 0.565 | 0.526 | 0.667 | 0.290 |
| | 5 | 0.000 | 0.000 | 0.715 | 0.956 | 0.737 | 0.897 | 0.692 |
| | 10 | 0.000 | 0.000 | 0.880 | 0.994 | 0.878 | 0.969 | 0.889 |
| | 25 | 0.000 | 0.000 | 0.995 | 1.000 | 0.991 | 1.000 | 1.000 |
| | 50 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 25 | 1 | 0.005 | 0.015 | 0.130 | 0.145 | 0.273 | 0.289 | 0.089 |
| | 5 | 0.000 | 0.009 | 0.193 | 0.252 | 0.382 | 0.424 | 0.157 |
| | 10 | 0.000 | 0.012 | 0.264 | 0.351 | 0.488 | 0.538 | 0.242 |
| | 25 | 0.000 | 0.006 | 0.444 | 0.585 | 0.743 | 0.803 | 0.415 |
| | 50 | 0.000 | 0.008 | 0.601 | 0.772 | 0.909 | 0.953 | 0.568 |
| 50 | 1 | 0.039 | 0.073 | 0.089 | 0.096 | 0.178 | 0.182 | 0.085 |
| | 5 | 0.017 | 0.062 | 0.099 | 0.108 | 0.211 | 0.225 | 0.104 |
| | 10 | 0.014 | 0.059 | 0.124 | 0.137 | 0.264 | 0.281 | 0.121 |
| | 25 | 0.011 | 0.076 | 0.163 | 0.192 | 0.383 | 0.403 | 0.145 |
| | 50 | 0.005 | 0.093 | 0.229 | 0.270 | 0.534 | 0.569 | 0.233 |
| 100 | 1 | 0.072 | 0.122 | 0.078 | 0.078 | 0.132 | 0.133 | 0.080 |
| | 5 | 0.052 | 0.100 | 0.080 | 0.081 | 0.148 | 0.149 | 0.073 |
| | 10 | 0.046 | 0.086 | 0.080 | 0.080 | 0.168 | 0.170 | 0.071 |
| | 25 | 0.068 | 0.142 | 0.106 | 0.110 | 0.235 | 0.237 | 0.087 |
| | 50 | 0.074 | 0.170 | 0.141 | 0.149 | 0.330 | 0.333 | 0.120 |
| 200 | 1 | 0.077 | 0.124 | 0.057 | 0.057 | 0.097 | 0.097 | 0.073 |
| | 5 | 0.084 | 0.118 | 0.071 | 0.071 | 0.051 | 0.050 | 0.066 |
| | 10 | 0.079 | 0.108 | 0.061 | 0.061 | 0.057 | 0.058 | 0.061 |
| | 25 | 0.102 | 0.132 | 0.080 | 0.079 | 0.047 | 0.047 | 0.068 |
| | 50 | 0.132 | 0.182 | 0.108 | 0.109 | 0.026 | 0.026 | 0.068 |

the short time series[3]. In addition, it is clear from Table 3.1 that when $T$ and $N$ increase, the empirical sizes of the panel-$t$ and standardized LR-bar tests approach the nominal size level of 5%, especially for $T = 200$ and $N \geq 5$. Furthermore, the empirical sizes of the group-$\rho$ and panel-$\rho$ tests are around the 5% level for $T = 100$, $N \geq 5$ and $T = 50$, $N \geq 5$, respectively. The size distortions of the group-$t$, panel-$t$ and standardized LR-bar tests decrease for fixed $N$ when $T$ grows.

### 3.4.2 Size-Adjusted Power Properties

The size-adjusted power results of the tests are demonstrated in Tables 3.2-3.7. Note that the case of $\psi_a, \psi_b, \psi_c \in \{0.5, 1\}$ and no correlation in the error terms is discussed only for $T = 10$. This is due to the fact that the powers of all the tests approach unity for $T \geq 25$ and $N \geq 10$. Thus, for $T = 10$, the standardized LR-bar and group-$t$ tests have the lowest power for the true cointegrating ranks of one and two (see Table 3.2). The panel-$\rho$ test has the highest power reaching 0.891 and 0.681 for the true cointegrating ranks of one and two, respectively. If the true cointegrating rank is three as in the lower part of Table 3.2, the power of rejecting the null hypothesis of no cointegration is the highest for the standardized LR-bar test (0.534), whereas the group-$t$ test has the lowest power (0.046).

Since the results when the $\psi$ parameters are 0.8 do not differ much from those in Table 3.2, they are not presented in detail to save space.

If the $\psi$ parameters are near unity, i.e. 0.95, and $T = 10$, the powers of all the tests are at most 0.074 for the true cointegrating ranks of one, two and three, which can be observed in Tables 3.3, 3.4 and 3.5. Table 3.3 indicates that the standardized LR-bar test has the lowest power and the panel-$\rho$ and panel-$t$ tests have the highest power. With the true cointegrating rank assumption of two, the last column of Table 3.4 indicates that the maximum-likelihood-based test has very low power for $T \leq 100$. In line with the theory, the powers of all the tests converge to unity for large $T$ and $N$. One interesting outcome of this Monte Carlo study is observed for $T = 100$ and the true cointegrating rank of three. Thus, as indicated in Table 3.5, the standardized LR-bar test has the highest power among all the tests. This is in contrast to the results in Tables 3.3 and 3.4.

[3]Hanck (2007, 2008) suggested to use the moments from the finite sample distribution. Moreover, he explained the increase of the size distortion with the increase in $N$ as the cumulative effect of small size distortions in the time series.

Table 3.2: Size-adjusted power results of the tests for $T = 10$ and $\theta_a = \theta_b = 0$.

| | N | group-$\rho$ | panel-$\rho$ | group-$t$(sd) | group-$t$(sc) | panel-$t$(sd) | panel-$t$(sc) | std. LR-bar |
|---|---|---|---|---|---|---|---|---|
| $\psi_a = 0.5,\ \psi_b = \psi_c = 1$ | 1 | 0.082 | 0.082 | 0.064 | 0.071 | 0.066 | 0.048 | 0.049 |
| | 5 | 0.188 | 0.204 | 0.072 | 0.091 | 0.102 | 0.103 | 0.047 |
| | 10 | 0.280 | 0.310 | 0.096 | 0.094 | 0.197 | 0.145 | 0.055 |
| | 25 | 0.494 | 0.613 | 0.098 | 0.129 | 0.436 | 0.382 | 0.055 |
| | 50 | 0.806 | 0.891 | 0.115 | 0.126 | 0.724 | 0.627 | 0.065 |
| $\psi_a = \psi_b = 0.5,\ \psi_c = 1$ | 1 | 0.069 | 0.069 | 0.058 | 0.064 | 0.055 | 0.056 | 0.045 |
| | 5 | 0.143 | 0.147 | 0.059 | 0.059 | 0.081 | 0.075 | 0.052 |
| | 10 | 0.213 | 0.233 | 0.095 | 0.099 | 0.141 | 0.098 | 0.090 |
| | 25 | 0.307 | 0.420 | 0.061 | 0.075 | 0.297 | 0.250 | 0.086 |
| | 50 | 0.587 | 0.681 | 0.066 | 0.082 | 0.462 | 0.361 | 0.134 |
| $\psi_a = \psi_b = \psi_c = 0.5$ | 1 | 0.061 | 0.061 | 0.048 | 0.044 | 0.043 | 0.038 | 0.073 |
| | 5 | 0.108 | 0.108 | 0.054 | 0.058 | 0.044 | 0.056 | 0.113 |
| | 10 | 0.141 | 0.154 | 0.073 | 0.058 | 0.089 | 0.055 | 0.181 |
| | 25 | 0.154 | 0.203 | 0.053 | 0.057 | 0.163 | 0.129 | 0.308 |
| | 50 | 0.291 | 0.385 | 0.039 | 0.046 | 0.255 | 0.189 | 0.534 |

**Table 3.3:** Size-adjusted power results of the tests for $\theta_a = \theta_b = 0$, $\psi_a = 0.95$ and $\psi_b = \psi_c = 1$.

| T | N | group-$\rho$ | panel-$\rho$ | group-$t$(sd) | group-$t$(sc) | panel-$t$(sd) | panel-$t$(sc) | std. LR-bar |
|---|---|---|---|---|---|---|---|---|
| 10 | 1 | 0.065 | 0.065 | 0.050 | 0.054 | 0.059 | 0.053 | 0.042 |
| | 5 | 0.051 | 0.054 | 0.055 | 0.054 | 0.047 | 0.048 | 0.039 |
| | 10 | 0.065 | 0.062 | 0.062 | 0.057 | 0.056 | 0.052 | 0.051 |
| | 25 | 0.056 | 0.062 | 0.051 | 0.056 | 0.070 | 0.065 | 0.042 |
| | 50 | 0.070 | 0.074 | 0.050 | 0.051 | 0.073 | 0.071 | 0.046 |
| 25 | 1 | 0.061 | 0.061 | 0.058 | 0.051 | 0.060 | 0.059 | 0.053 |
| | 5 | 0.078 | 0.088 | 0.066 | 0.058 | 0.074 | 0.079 | 0.045 |
| | 10 | 0.084 | 0.087 | 0.073 | 0.079 | 0.102 | 0.108 | 0.031 |
| | 25 | 0.103 | 0.160 | 0.078 | 0.092 | 0.151 | 0.146 | 0.041 |
| | 50 | 0.136 | 0.194 | 0.095 | 0.110 | 0.221 | 0.227 | 0.017 |
| 50 | 1 | 0.052 | 0.052 | 0.056 | 0.061 | 0.056 | 0.057 | 0.038 |
| | 5 | 0.129 | 0.140 | 0.105 | 0.109 | 0.121 | 0.114 | 0.046 |
| | 10 | 0.141 | 0.192 | 0.111 | 0.110 | 0.185 | 0.185 | 0.045 |
| | 25 | 0.230 | 0.305 | 0.195 | 0.204 | 0.333 | 0.334 | 0.034 |
| | 50 | 0.395 | 0.561 | 0.324 | 0.321 | 0.557 | 0.562 | 0.034 |
| 100 | 1 | 0.093 | 0.093 | 0.076 | 0.076 | 0.081 | 0.082 | 0.067 |
| | 5 | 0.250 | 0.282 | 0.183 | 0.182 | 0.230 | 0.230 | 0.076 |
| | 10 | 0.436 | 0.473 | 0.314 | 0.316 | 0.437 | 0.438 | 0.103 |
| | 25 | 0.715 | 0.844 | 0.560 | 0.565 | 0.817 | 0.812 | 0.141 |
| | 50 | 0.940 | 0.988 | 0.855 | 0.860 | 0.975 | 0.974 | 0.169 |
| 200 | 1 | 0.199 | 0.199 | 0.165 | 0.165 | 0.215 | 0.214 | 0.102 |
| | 5 | 0.680 | 0.753 | 0.654 | 0.658 | 0.866 | 0.868 | 0.279 |
| | 10 | 0.962 | 0.981 | 0.935 | 0.935 | 0.991 | 0.989 | 0.454 |
| | 25 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.782 |
| | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.977 |

**Table 3.4:** Size-adjusted power results of the tests for $\theta_a = \theta_b = 0$, $\psi_a = \psi_b = 0.95$ and $\psi_c = 1$.

| T | N | group-$\rho$ | panel-$\rho$ | group-$t$(sd) | group-$t$(sc) | panel-$t$(sd) | panel-$t$(sc) | std. LR-bar |
|---|---|---|---|---|---|---|---|---|
| 10 | 1 | 0.058 | 0.058 | 0.044 | 0.060 | 0.051 | 0.054 | 0.043 |
| | 5 | 0.044 | 0.047 | 0.049 | 0.048 | 0.042 | 0.051 | 0.041 |
| | 10 | 0.062 | 0.058 | 0.055 | 0.052 | 0.060 | 0.042 | 0.045 |
| | 25 | 0.053 | 0.059 | 0.049 | 0.058 | 0.060 | 0.056 | 0.035 |
| | 50 | 0.062 | 0.071 | 0.042 | 0.044 | 0.067 | 0.062 | 0.041 |
| 25 | 1 | 0.065 | 0.065 | 0.058 | 0.045 | 0.056 | 0.055 | 0.048 |
| | 5 | 0.062 | 0.074 | 0.047 | 0.045 | 0.062 | 0.070 | 0.038 |
| | 10 | 0.071 | 0.078 | 0.059 | 0.061 | 0.097 | 0.090 | 0.029 |
| | 25 | 0.073 | 0.107 | 0.068 | 0.080 | 0.135 | 0.130 | 0.020 |
| | 50 | 0.104 | 0.130 | 0.079 | 0.094 | 0.188 | 0.189 | 0.013 |
| 50 | 1 | 0.055 | 0.055 | 0.053 | 0.060 | 0.058 | 0.062 | 0.043 |
| | 5 | 0.103 | 0.116 | 0.095 | 0.096 | 0.111 | 0.106 | 0.049 |
| | 10 | 0.113 | 0.157 | 0.098 | 0.095 | 0.163 | 0.170 | 0.044 |
| | 25 | 0.185 | 0.241 | 0.164 | 0.173 | 0.298 | 0.309 | 0.049 |
| | 50 | 0.288 | 0.422 | 0.267 | 0.261 | 0.470 | 0.475 | 0.040 |
| 100 | 1 | 0.074 | 0.074 | 0.064 | 0.066 | 0.072 | 0.075 | 0.063 |
| | 5 | 0.213 | 0.236 | 0.156 | 0.156 | 0.216 | 0.216 | 0.157 |
| | 10 | 0.348 | 0.384 | 0.277 | 0.276 | 0.391 | 0.393 | 0.235 |
| | 25 | 0.616 | 0.749 | 0.489 | 0.489 | 0.758 | 0.758 | 0.418 |
| | 50 | 0.858 | 0.960 | 0.783 | 0.785 | 0.953 | 0.951 | 0.640 |
| 200 | 1 | 0.165 | 0.165 | 0.134 | 0.134 | 0.191 | 0.191 | 0.231 |
| | 5 | 0.614 | 0.707 | 0.608 | 0.615 | 0.845 | 0.846 | 0.825 |
| | 10 | 0.928 | 0.966 | 0.915 | 0.912 | 0.98 | 0.98 | 0.973 |
| | 25 | 1.000 | 1.000 | 0.999 | 0.999 | 1.000 | 1.000 | 1.000 |
| | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

**Table 3.5:** Power results of the tests for $\theta_a = \theta_b = 0$, $\psi_a = \psi_b = \psi_c = 0.95$.

| T | N | group-$\rho$ | panel-$\rho$ | group-$t$(sd) | group-$t$(sc) | panel-$t$(sd) | panel-$t$(sc) | std. LR-bar |
|---|---|---|---|---|---|---|---|---|
| 10 | 1 | 0.057 | 0.057 | 0.044 | 0.053 | 0.054 | 0.051 | 0.052 |
| | 5 | 0.056 | 0.050 | 0.040 | 0.043 | 0.042 | 0.052 | 0.036 |
| | 10 | 0.054 | 0.045 | 0.070 | 0.069 | 0.059 | 0.040 | 0.046 |
| | 25 | 0.042 | 0.043 | 0.050 | 0.057 | 0.060 | 0.054 | 0.042 |
| | 50 | 0.058 | 0.054 | 0.046 | 0.038 | 0.052 | 0.046 | 0.034 |
| 25 | 1 | 0.059 | 0.059 | 0.057 | 0.049 | 0.056 | 0.050 | 0.064 |
| | 5 | 0.053 | 0.064 | 0.041 | 0.047 | 0.059 | 0.070 | 0.042 |
| | 10 | 0.047 | 0.054 | 0.049 | 0.050 | 0.077 | 0.079 | 0.029 |
| | 25 | 0.060 | 0.090 | 0.058 | 0.064 | 0.105 | 0.108 | 0.025 |
| | 50 | 0.077 | 0.095 | 0.069 | 0.072 | 0.155 | 0.159 | 0.021 |
| 50 | 1 | 0.051 | 0.051 | 0.057 | 0.059 | 0.068 | 0.071 | 0.062 |
| | 5 | 0.083 | 0.084 | 0.075 | 0.076 | 0.098 | 0.093 | 0.066 |
| | 10 | 0.091 | 0.117 | 0.083 | 0.080 | 0.143 | 0.134 | 0.085 |
| | 25 | 0.118 | 0.156 | 0.113 | 0.124 | 0.233 | 0.238 | 0.103 |
| | 50 | 0.186 | 0.278 | 0.191 | 0.191 | 0.376 | 0.379 | 0.157 |
| 100 | 1 | 0.073 | 0.073 | 0.057 | 0.056 | 0.070 | 0.073 | 0.162 |
| | 5 | 0.173 | 0.189 | 0.128 | 0.129 | 0.192 | 0.193 | 0.375 |
| | 10 | 0.272 | 0.306 | 0.229 | 0.223 | 0.332 | 0.331 | 0.599 |
| | 25 | 0.499 | 0.625 | 0.414 | 0.409 | 0.686 | 0.690 | 0.936 |
| | 50 | 0.751 | 0.904 | 0.699 | 0.702 | 0.927 | 0.927 | 0.997 |
| 200 | 1 | 0.161 | 0.161 | 0.146 | 0.147 | 0.183 | 0.185 | 0.536 |
| | 5 | 0.573 | 0.646 | 0.571 | 0.575 | 0.823 | 0.830 | 0.999 |
| | 10 | 0.889 | 0.954 | 0.886 | 0.885 | 0.975 | 0.975 | 1.000 |
| | 25 | 0.999 | 1.000 | 0.998 | 0.998 | 1.000 | 1.000 | 1.000 |
| | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

**Table 3.6:** Size-adjusted power results of the tests for $\theta_a = \theta_b = 0.7$, $\psi_a = 0.95$ and $\psi_b = \psi_c = 1$.

| T | N | group-$\rho$ | panel-$\rho$ | group-$t$(sd) | group-$t$(sc) | panel-$t$(sd) | panel-$t$(sc) | std. LR-bar |
|---|---|---|---|---|---|---|---|---|
| 10 | 1 | 0.043 | 0.043 | 0.045 | 0.042 | 0.511 | 0.411 | 0.188 |
| | 5 | 0.020 | 0.011 | 0.038 | 0.034 | 0.930 | 0.892 | 0.428 |
| | 10 | 0.014 | 0.006 | 0.032 | 0.031 | 0.996 | 0.988 | 0.647 |
| | 25 | 0.003 | 0.001 | 0.028 | 0.029 | 1.000 | 1.000 | 0.923 |
| | 50 | 0.000 | 0.000 | 0.024 | 0.024 | 1.000 | 1.000 | 0.996 |
| 25 | 1 | 0.026 | 0.026 | 0.031 | 0.026 | 0.676 | 0.666 | 0.653 |
| | 5 | 0.007 | 0.002 | 0.009 | 0.012 | 0.943 | 0.941 | 0.971 |
| | 10 | 0.001 | 0.000 | 0.000 | 0.001 | 0.997 | 0.995 | 0.998 |
| | 25 | 0.002 | 0.000 | 0.001 | 0.001 | 1.000 | 1.000 | 1.000 |
| | 50 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 |
| 50 | 1 | 0.016 | 0.016 | 0.031 | 0.030 | 0.735 | 0.735 | 0.916 |
| | 5 | 0.003 | 0.003 | 0.002 | 0.003 | 0.970 | 0.970 | 1.000 |
| | 10 | 0.000 | 0.000 | 0.000 | 0.000 | 0.998 | 0.999 | 1.000 |
| | 25 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 |
| | 50 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 |
| 100 | 1 | 0.016 | 0.016 | 0.015 | 0.016 | 0.746 | 0.749 | 1.000 |
| | 5 | 0.003 | 0.011 | 0.003 | 0.002 | 0.991 | 0.991 | 1.000 |
| | 10 | 0.004 | 0.007 | 0.001 | 0.001 | 1.000 | 1.000 | 1.000 |
| | 25 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 |
| | 50 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 |
| 200 | 1 | 0.048 | 0.048 | 0.040 | 0.041 | 0.881 | 0.883 | 1.000 |
| | 5 | 0.087 | 0.134 | 0.084 | 0.088 | 0.999 | 0.999 | 1.000 |
| | 10 | 0.112 | 0.223 | 0.091 | 0.088 | 1.000 | 1.000 | 1.000 |
| | 25 | 0.158 | 0.417 | 0.126 | 0.127 | 1.000 | 1.000 | 1.000 |
| | 50 | 0.216 | 0.784 | 0.178 | 0.178 | 1.000 | 1.000 | 1.000 |

**Table 3.7:** Size-adjusted power results of the tests for $\theta_a = \theta_b = 0.7$, $\psi_a = \psi_b = 0.95$ and $\psi_c = 1$.

| T | N | group-$\rho$ | panel-$\rho$ | group-$t$(sd) | group-$t$(sc) | panel-$t$(sd) | panel-$t$(sc) | std. LR-bar |
|---|---|---|---|---|---|---|---|---|
| 10 | 1 | 0.041 | 0.041 | 0.044 | 0.043 | 0.337 | 0.288 | 0.212 |
| | 5 | 0.030 | 0.017 | 0.033 | 0.031 | 0.779 | 0.688 | 0.435 |
| | 10 | 0.011 | 0.007 | 0.040 | 0.037 | 0.946 | 0.879 | 0.646 |
| | 25 | 0.003 | 0.000 | 0.032 | 0.031 | 0.999 | 0.999 | 0.919 |
| | 50 | 0.000 | 0.000 | 0.031 | 0.026 | 1.000 | 1.000 | 0.991 |
| 25 | 1 | 0.033 | 0.033 | 0.043 | 0.037 | 0.514 | 0.501 | 0.635 |
| | 5 | 0.009 | 0.006 | 0.012 | 0.009 | 0.821 | 0.832 | 0.975 |
| | 10 | 0.001 | 0.000 | 0.003 | 0.004 | 0.951 | 0.952 | 1.000 |
| | 25 | 0.000 | 0.000 | 0.001 | 0.002 | 0.999 | 0.999 | 1.000 |
| | 50 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 |
| 50 | 1 | 0.029 | 0.029 | 0.033 | 0.034 | 0.585 | 0.586 | 0.926 |
| | 5 | 0.007 | 0.006 | 0.013 | 0.011 | 0.890 | 0.890 | 1.000 |
| | 10 | 0.004 | 0.002 | 0.002 | 0.003 | 0.983 | 0.986 | 1.000 |
| | 25 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 |
| | 50 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 |
| 100 | 1 | 0.026 | 0.026 | 0.028 | 0.028 | 0.596 | 0.605 | 1.000 |
| | 5 | 0.019 | 0.024 | 0.021 | 0.022 | 0.956 | 0.956 | 1.000 |
| | 10 | 0.011 | 0.016 | 0.009 | 0.008 | 0.999 | 0.999 | 1.000 |
| | 25 | 0.003 | 0.004 | 0.002 | 0.001 | 1.000 | 1.000 | 1.000 |
| | 50 | 0.000 | 0.005 | 0.000 | 0.000 | 1.000 | 1.000 | 1.000 |
| 200 | 1 | 0.086 | 0.086 | 0.070 | 0.070 | 0.761 | 0.763 | 1.000 |
| | 5 | 0.238 | 0.282 | 0.245 | 0.249 | 0.998 | 0.998 | 1.000 |
| | 10 | 0.443 | 0.558 | 0.390 | 0.382 | 1.000 | 1.000 | 1.000 |
| | 25 | 0.730 | 0.888 | 0.710 | 0.712 | 1.000 | 1.000 | 1.000 |
| | 50 | 0.942 | 0.997 | 0.945 | 0.946 | 1.000 | 1.000 | 1.000 |

Tables 3.6 and 3.7 present the results on how the tests behave under the assumption of correlated error terms. Only the case with the highest correlation parameters, i.e. $\theta_a = \theta_b = 0.7$, is discussed here as the power results do not change drastically if the correlation between the error terms is low. The same is true for the $\psi$ parameters. For $\psi_a, \psi_b, \psi_c \in \{0.95, 1\}$ the powers of the standardized LR-bar and panel-$t$ tests approach unity, even if $T = 10$. In contrast, the powers of the other tests are near zero for $T \leq 100$ (see the upper parts of Tables 3.6 and 3.7). The power of rejecting the cointegrating rank of zero for the group-$\rho$ and group-$t$ tests does not go to unity for the true cointegrating rank of one, even if $T$ and $N$ are large. The power of all the tests converge to unity if the true cointegrating rank is two (see Table 3.7). Moreover, the powers of the group-$t$ and panel-$t$ tests are not much different with the increase in $T$ and $N$ when either the Schwarz Criterion or the step-down lag selection method is used.

## 3.5 Empirical Example: Fisher Hypothesis

In this section, I use the Fisher hypothesis as an empirical example to demonstrate to which extend the panel cointegration analysis gives different results than the results of the standard cointegration testing techniques.

In empirical literature, there are controversial conclusions on the existence of the Fisher hypothesis. The nonstationarity of the nominal interest rate and evidence on possible nonstationarity of the inflation rate make the application of the cointegration testing techniques a reasonable choice to test for the long-run relation between the nominal interest rate and the inflation rate. Relevant studies which find evidence for Fisher relation using the unit root and cointegration testing techniques are: Atkins (1989), Evans and Lewis (1995), Crowder and Hoffman (1996), Crowder (1997). In contrast to this, the studies of Rose (1988), MacDonald and Murphy (1989), Mishkin (1992) and Dutt and Gosh (1995) cannot find any evidence for the Fisher effect. A panel data study by Crowder (2003) with 9 industrialized countries concludes that the Fisher effect exists.

The Fisher hypothesis states that the real interest rate ($r_{it}$) is the difference between the nominal interest rate ($n_{it}$) and the expected inflation rate ($\pi_{it}^e$).

$$r_{it} = n_{it} - \pi_{it}^e, \tag{3.8}$$

which means that nobody lends at a nominal rate lower than the expected inflation rate, and the nominal interest rate is equal to the cost of borrowing plus the expected inflation rate.

Furthermore, Fisher (1930) stated that the real interest rate is constant or shows little trend in the long-run. This can be explained by the phenomenon that the nominal interest rate absorbs all the changes in the expected inflation rate. If the real interest rate changes with a change in the expected inflation, then the Fisher hypothesis does not hold. Assuming stationarity of the real interest rate around a positive constant ($r^*$) and a normally distributed error term ($u_{it} \sim N(0, \sigma_{iu}^2)$), (3.8) becomes

$$r_{it} = r^* + u_{it}. \tag{3.9}$$

Additionally, the actual inflation rate ($\pi_{it}$) differs from the expected inflation rate ($\pi_{it}^e$) by a random stationary error term ($\xi_{it} \sim N(0, \sigma_{i\xi}^2)$), i.e. the agents do not make systematic errors. Thus,

$$\pi_{it} = \pi_{it}^e + \xi_{it}. \tag{3.10}$$

On account of (3.9), (3.10) and $n_{it} = r_{it} + \pi_{it}^e$, the Fisher equation used in the cointegration analysis is written as

$$n_{it} = a + b\pi_{it} + \varepsilon_{it}, \tag{3.11}$$

in which $a = r^*$, $\varepsilon_{it} = u_{it} - \xi_{it}$ and according to theory $b = 1$. In order to see if the Fisher relation holds, I search for the existence of a cointegrating relation between the nominal interest rate and the inflation rate.

To test the Fisher hypothesis two different datasets consisting of the three-month nominal interest rate and the three-month inflation rate are considered. The first dataset, called dataset A, consists of monthly data for 19 OECD countries[4] from 1989:06 to 1998:12 (i.e. $T = 115$) and the second dataset, called dataset B, consists of monthly data for 11 OECD countries[5] from 1991:02 to 2002:12 (i.e. $T = 145$). A detailed description of the variables and the sources of the data can be found in the Appendix B.1.

The results of the different panel cointegration tests are presented in Tables 3.8-3.9. The standardized LR-bar statistic is not included in the analysis as the test does not allow for deterministic terms.

To standardize the test statistics of Pedroni (1999), the asymptotic mean and variance values for the model with one independent variable ($K = 1$) are required, which can be found in Pedroni (1995). For the ADF statistic-based panel cointegration tests of Pedroni, two lag selection methods are considered:

[4] Austria, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Japan, Mexico, Netherlands, Norway, Portugal, Spain, Sweden, UK, US.

[5] Canada, Denmark, Hungary, Iceland, Japan, Korea, Mexico, Norway, Sweden, UK, US.

The step-down method and the Schwarz criterion. The maximum lag order is limited to 12 because the datasets consist of monthly data.

Note again, with the residual-based panel cointegration tests one cannot test for the rank of the cointegrating matrix. It can just be tested whether there is a cointegrating relation or not.

Some of the residual-based panel cointegration tests give different results for both datasets[6]. For dataset A only the panel-$\rho$ test cannot reject the null hypothesis of no cointegration which means that the Fisher hypothesis does not hold (see Table 3.8). Moreover, all the tests point out the existence of a cointegrating relation between the nominal interest rate and the inflation rate for dataset B at the 5% significance level (see Table 3.9).

**Table 3.8:** Results of the panel cointegration tests of Pedroni for dataset A.

| Group Tests | | Panel Tests | |
|---|---|---|---|
| Group-$\rho$ | -3.23* | Panel-$\rho$ | -1.50 |
| Group-$t$(sc) | -7.53* | Panel-$t$(sc) | -6.10* |
| Group-$t$(sd) | -4.50* | Panel-$t$(sd) | -5.00* |

*Notes: * indicates rejection of the null hypothesis of no cointegration.*

**Table 3.9:** Results of the panel cointegration tests of Pedroni for dataset B.

| Group Tests | | Panel Tests | |
|---|---|---|---|
| Group-$\rho$ | -3.84* | Panel-$\rho$ | -2.22* |
| Group-$t$(sc) | -7.32* | Panel-$t$(sc) | -2.98* |
| Group-$t$(sd) | -5.24* | Panel-$t$(sd) | -2.29* |

*Notes: * indicates rejection of the null hypothesis of no cointegration.*

## 3.6 Conclusions

With the extensive simulation study in Section 3.4 it can be concluded that the panel-$t$ test has the best size and size-adjusted power properties. The size-adjusted power of the panel-$t$ test approaches unity for small $T$ and $N$, even if there is strong correlation between the innovations to the stationary and nonstationary components of the DGP, whereas its empirical size is around the nominal 5% level when $T = 200$ and $N \geq 5$. On the contrary, the other three residual-based panel cointegration tests (group-$\rho$, panel-$\rho$ and group-$t$

[6]At the 5% level the residual-based tests of Pedroni have a critical value of -1.65.

tests) have poor size-adjusted powers if the correlation and $\psi$ parameters are high (e.g. for $\theta_a = \theta_b = 0.7$ and $\psi_a = \psi_b = 0.95$, respectively).

The second best test, in terms of size and size-adjusted power properties is the standardized LR-bar test. It has better size-adjusted power than the other tests if the correlation parameter is high and the $\psi$ parameter are around unity. The empirical size of the standardized LR-bar test is around the 5% level, similar to the size of the panel-$t$ test, especially if the time dimension increases faster than the cross-section dimension as theory points out. The standardized LR-bar test has also high size-adjusted power, mainly for large $T$.

It should also be emphasized that the size and size-adjusted power results of the residual-based panel cointegration tests can depend on the variable chosen to normalize the cointegrating relation. In this study for the residual-based panel cointegration tests, the first variable of the DGP is used to normalize the cointegrating relation. On the contrary, the maximum-likelihood based tests do not depend on the variable used for the normalization of the cointegrating relation.

In Section 3.5 I empirically analyzed the validity of the Fisher hypothesis using residual-based panel cointegration tests of Pedroni (1999). The tests pointed out the existence of the a long-run relation between the nominal interest rate and the inflation rate for two different datasets consisting of OECD countries.

# Chapter 4

# Existence of the Moments of the Asymptotic Trace Statistic

In this chapter the incorrectness of the proof of Lemma 1 in Larsson et al. (2001) related to the existence and finiteness of the moments of the asymptotic trace statistic is shown and explained in detail. Moreover, the proof is corrected for the case when the difference between the number of variables in the system ($K$) and the number of cointegrating relations ($r$) is one, i.e. $d = K - r = 1$.

Larsson et al. (2001) proposed the standardized LR-bar test statistic for panel data, explained in Chapter 2, which is based on the mean of the individual rank trace statistic of Johansen (1995). Under the null hypothesis this panel cointegration test statistic is asymptotically $N(0, 1)$ distributed as $T \to \infty$ and $N \to \infty$, in such a way that $\sqrt{N}/T \to 0$. In order to standardize the LR-bar statistic, the first two moments of the asymptotic trace statistic must exist and be finite. This is necessary to establish the asymptotic distribution of the standardized LR-bar test statistic.

The outline of this chapter is as follows: In Section 4.1 it is demonstrated that the proof of Lemma 1 in Larsson et al. (2001) is not correct. Section 4.2 shows that the upper bound of the first moments of the trace statistic depends on the number of time observations. Section 4.3 gives the corrected proof on the finiteness of the moments of the asymptotic trace statistic ($Z_d$) for $d = 1$, with $d$ being the dimension of the Brownian motion functionals underlying $Z_d$. A final conclusion is provided in Section 4.4.

## 4.1 Comments on the Proof of Lemma 1 in Larsson et al. (2001)

Lemma 1 in Larsson et al. (2001) says that the second moment $E(Z_d^2)$ exists and is finite. The asymptotic trace statistic $Z_d$ is defined as

$$Z_d = \mathrm{tr}\left[\int_0^1 \mathrm{d}W(s)W(s)'\left(\int_0^1 W(s)W(s)'\mathrm{d}s\right)^{-1}\int_0^1 W(s)\mathrm{d}W(s)'\right], \quad (4.1)$$

in which $W(s)$ is a $d = (K-r)$-dimensional Brownian motion. Larsson et al. (2001) based their proof of Lemma 1 on the uniform boundedness of the first four moments of the trace statistic $(Z_{T,d})$ implying the uniform integrability of $Z_{T,d}^2$.

To prove Lemma 1, Larsson et al. (2001) aimed to show that $E(Z_{T,d}^2) < a$ for all $T > d+8$, with some finite and positive constant $a$. They used the fact that as $T \to \infty$

$$Z_{T,d} = \mathrm{tr}\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t X_{t-1}'\left(\frac{1}{T^2}\sum_{t=1}^{T}X_{t-1}X_{t-1}'\right)^{-1}\frac{1}{T}\sum_{t=1}^{T}X_{t-1}\varepsilon_t'\right] \xrightarrow{w} Z_d, \quad (4.2)$$

in which $Z_{T,d}$ is the trace statistic and $\varepsilon_t \sim N_d(0,\Omega)$ *i.i.d.* Larsson et al. (2001) defined $(T \times d)$ matrices $\varepsilon = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_T)'$, $X = (X_1, X_2, ..., X_T)'$, with $X_t = \sum_{i=1}^{t}\varepsilon_i$ for $t = 1, \ldots, T$ and $(T \times T)$ matrices $A$ and $B$.

$$A \equiv \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & \cdots & \cdots & \cdots & 1 \end{pmatrix}, \qquad B \equiv \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Using the matrix $A$, the matrix $X$ can also be formulated as $X = A\varepsilon$, so that the sums in (4.2) can be represented by

$$\sum_{t=1}^{T}\varepsilon_t X_{t-1}' = \varepsilon' BA\varepsilon; \ \sum_{t=1}^{T}X_{t-1}\varepsilon_t' = \varepsilon' A'B'\varepsilon; \ \sum_{t=1}^{T}X_{t-1}X_{t-1}' = \varepsilon' A'B'BA\varepsilon. \quad (4.3)$$

Additionally, Larsson et al. (2001) introduced $C = T^{-1}BA\varepsilon$, and they obtained

$$\frac{1}{T^2}\sum_{t=1}^{T}X_{t-1}X_{t-1}' = T^{-2}\varepsilon' A'B'BA\varepsilon = C'C = G'\Lambda G, \quad (4.4)$$

in which $\Lambda = \text{diag}(\lambda_1, ..., \lambda_d)$ is the matrix of the eigenvalues of the symmetric positive definite matrix $C'C$ and $G$ is an orthogonal $(d \times d)$ matrix of eigenvectors corresponding to the eigenvalues of $C'C$.

The first comment about the proof of Larsson et al. (2001) is related to the $G$ matrix. In their proof Larsson et al. (2001) neglected that $G$ is a random matrix and they proceeded under the assumption that $G$ is a deterministic matrix. However, $G$ is indeed a random matrix because it depends on the random variable $\varepsilon_t$.

Neglecting the randomness of $G$, Larsson et al. (2001) defined $\varepsilon = \tilde{\varepsilon}G$. They rewrote (4.1) using (4.3) and $\varepsilon = \tilde{\varepsilon}G$, which delivers the following expression.

$$\begin{aligned} Z_{T,d} &= T^{-2}\text{tr}(\varepsilon' BA\varepsilon G'\Lambda^{-1}G\varepsilon' A'B'\varepsilon) = T^{-2}\text{tr}(\tilde{\varepsilon}' BA\tilde{\varepsilon}\Lambda^{-1}\tilde{\varepsilon}' A'B'\tilde{\varepsilon}) \\ &= \text{tr}(H\Lambda^{-1}) = \sum_{i=1}^{d} H_{ii}\lambda_i^{-1}, \qquad (4.5) \end{aligned}$$

with $H = T^{-2}\tilde{\varepsilon}' A'B'\tilde{\varepsilon}\tilde{\varepsilon}' BA\tilde{\varepsilon}$ and $\Lambda = T^{-2}\tilde{\varepsilon}' A'B'BA\tilde{\varepsilon}$. Larsson et al. (2001) stated that $H$ has the same distribution as $\varepsilon$ because of the orthogonality of $G$. As a result the elements of $H$ are the sum of products of $\chi^2$ and normal variables. In addition to this, Larsson et al. (2001) defined $c_n(T) \equiv \max_{1\le i,j \le d} E|H_{ij}^n|$. Using the triangle and Cauchy-Schwarz inequalities they demonstrated that

$$\begin{aligned} E(Z_{T,d}^2) &= E\left(\left|\sum_{i=1}^{d}\sum_{j=1}^{d} H_{ii}\lambda_i^{-1}H_{jj}\lambda_j^{-1}\right|\right) \qquad (4.6) \\ &\le \sum_{i=1}^{d}\sum_{j=1}^{d}[E(H_{ii}^4)E(\lambda_i^{-4})E(H_{jj}^4)E(\lambda_j^{-4})]^{\frac{1}{4}}. \qquad (4.7) \end{aligned}$$

Since all $E(H_{ii}^4) \le c_4$, for the finiteness of $E(Z_{T,d}^2)$ Larsson et al. (2001) should have exhibited that $E(\lambda_i^{-4})$ are finite for all $i$. Hence, they stated that $\Lambda = T^{-2}G\varepsilon' A'B'BA\varepsilon G' = T^{-2}\tilde{\varepsilon}' A'B'BA\tilde{\varepsilon}$ is Wishart distributed with parameters $d$, $T-1$ and $\Sigma_T$, i.e $\Lambda \sim W_d(T-1, \Sigma_T)$. If $\Lambda$ is Wishart distributed, all its diagonal elements will have finite fourth inverse moments. However, according to the definition of Muirhead (1982)[1], $\Lambda$ cannot be Wishart distributed because $G$ is dependent on the matrix $\varepsilon$. The Wishart distribution requires that the rows of $BA\varepsilon G'$ are independent and normally distributed

[1]If $X = Y'Y$, in which the $(n \times m)$ matrix $Y$ is $N(0, I_n \otimes \Sigma)$, then $X$ follows a *Wishart distribution* with $n$ degrees of freedom and covariance matrix $\Sigma$: i.e. $X \sim W_m(n, \Sigma)$ where $m$ denotes the size of the matrix X. Note that $I_n \otimes \Sigma$ is the covariance matrix of $\mathbf{y} = \text{vec}(Y')$.

with some common covariance matrix, but this is not valid here because $G$ depends on $\varepsilon$.

$\Lambda$ can be Wishart distributed[2] if $A'B'BA$ is a symmetric and idempotent matrix and $\widetilde{\varepsilon} \sim N(0, I_T \otimes \Omega)$, which is also not fulfilled. Moreover, $\Lambda$ is a diagonal matrix and a diagonal matrix cannot be Wishart distributed. Thus, the statement about the Wishart distribution of $\Lambda$ is wrong. The finiteness of the first two moments of the asymptotic trace statistic $Z_d$ should be proved using another method.

## 4.2 On the Finiteness of the Moments of the Trace Statistic

This section demonstrates that an upper bound for the first moments of $Z_{T,d}$ depends on the number of time observations $T$, and tends to infinity as $T \to \infty$. In other words, the upper bound for $E(Z_{T,d})$ is infinite.

To prove the finiteness of the moments of the trace statistic, let us use the following expressions.

$$A_T = \frac{1}{T^2}\varepsilon' A'B'BA\varepsilon, \quad B_T = \frac{1}{T}\varepsilon' A'B'\varepsilon. \tag{4.8}$$

Note that $\varepsilon$, $X$, $A$ and $B$ are the same matrices as in Section 4.1.

Using the $(d \times d)$ matrices $A_T$ and $B_T$, the expression (4.2) can be rewritten as

$$\begin{aligned} Z_{T,d} &= \operatorname{tr}(B_T' A_T^{-1} B_T) = \operatorname{tr}\left[\varepsilon' BA\varepsilon(\varepsilon' A'B'BA\varepsilon)^{-1}\varepsilon' A'B'\varepsilon\right] & (4.9)\\ &= \operatorname{tr}\left[\varepsilon' D\varepsilon(\varepsilon' D'D\varepsilon)^{-1}\varepsilon' D'\varepsilon\right]. & (4.10)\end{aligned}$$

with $D = BA$. Let $Y = D\varepsilon$ and

$$P_Y = Y(Y'Y)^{-1}Y' = D\varepsilon(\varepsilon' D'D\varepsilon)^{-1}\varepsilon' D' \tag{4.11}$$

be the projection matrix onto the column space of $Y$. Therefore,

$$Z_{T,d} = \operatorname{tr}[\varepsilon' D\varepsilon(\varepsilon' D'D\varepsilon)^{-1}\varepsilon' D'\varepsilon] = \operatorname{tr}(\varepsilon' P_Y \varepsilon) \leq \operatorname{tr}(\varepsilon'\varepsilon) \tag{4.12}$$

because $(I_T - P_Y)$ is a nonnegative definite matrix. Consequently, $\varepsilon'(I_T - P_Y)\varepsilon = \varepsilon'\varepsilon - \varepsilon' P_Y \varepsilon$ is also a nonnegative definite matrix.

[2]see Muirhead, 1982: Let $X$ be an $(n \times m)$ random matrix and $P$ be an $(n \times n)$ symmetric and idempotent matrix of rank $k \geq m$. If $X$ is $N(0, I_n \otimes \Sigma)$, then $X'PX$ is $W_m(k, \Sigma)$.

From the definition of the Wishart distribution we can conclude that $\varepsilon'\varepsilon$ is Wishart distributed with $T$ degrees of freedom and covariance matrix $\Omega$ as $\varepsilon \sim N_d(0, I_T \otimes \Omega)$. This can be denoted by $\varepsilon'\varepsilon \sim W_d(T, \Omega)$. Note that $\varepsilon'\varepsilon$ has a density function[3] for $T \geq d$, i.e. if $d = 1$ the density function exists for all $T$. Additionally, all the moments of $Z_{T,d}$ exist as all the moments of the Wishart distributed $\varepsilon'\varepsilon$ exist (see Letac and Massam, 1999; Graczyk et al., 2005).

To prove the finiteness of the expression $E(Z_{T,d})$, we should compute $E[\mathrm{tr}(\varepsilon'\varepsilon)]$ as

$$E(Z_{T,d}) = E[\mathrm{tr}(\varepsilon' P \varepsilon)] \leq E[\mathrm{tr}(\varepsilon'\varepsilon)]. \tag{4.13}$$

Let the $d$-dimensional column vector $\varepsilon_t$ be $\varepsilon_t = (\varepsilon_{t1}, \ldots, \varepsilon_{td})' \sim N(0, \Omega)$ *i.i.d.* Then,

$$\varepsilon'\varepsilon = \sum_{t=1}^{T} \varepsilon_t \varepsilon_t' \text{ with } \varepsilon_t \varepsilon_t' = \begin{pmatrix} \varepsilon_{t1} \\ \vdots \\ \varepsilon_{td} \end{pmatrix} (\varepsilon_{t1}, \ldots, \varepsilon_{td}) = [\varepsilon_{ti}\varepsilon_{tj}]_{i,j=1,..,d}. \tag{4.14}$$

Thus,

$$\mathrm{tr}(\varepsilon_t \varepsilon_t') = \sum_{i=1}^{d} \varepsilon_{ti}^2 \quad \text{with} \quad \varepsilon_{ti} \sim N(0, \omega_i^2), \quad \omega_i^2 = \omega_{ii}. \tag{4.15}$$

Using the relation $\varepsilon_{ti}^2 = \omega_i^2 Z_{ti}$ and $Z_{ti} \sim \chi^2_{(1)}$, we can obtain the expression for $\mathrm{tr}(\varepsilon_t \varepsilon_t')$.

$$\mathrm{tr}(\varepsilon_t \varepsilon_t') = \sum_{i=1}^{d} \omega_i^2 Z_{ti} \tag{4.16}$$

This also leads to the relation

$$\mathrm{tr}(\varepsilon'\varepsilon) = \sum_{t=1}^{T} \mathrm{tr}(\varepsilon_t \varepsilon_t') = \sum_{t=1}^{T} \sum_{i=1}^{d} \omega_i^2 Z_{ti} = \sum_{i=1}^{d} \omega_i^2 \sum_{t=1}^{T} Z_{ti} = \sum_{i=1}^{d} \omega_i^2 \xi_i, \tag{4.17}$$

in which $\sum_{t=1}^{T} Z_{ti} = \xi_i \sim \chi^2_{(T)}$. Hence,

$$E(Z_{T,d}) \quad \leq \quad E[\mathrm{tr}(\varepsilon'\varepsilon)] \leq E\left(\sum_{i=1}^{d} \omega_i^2 \xi_i\right) = \sum_{i=1}^{d} \omega_i^2 \underbrace{E(\xi_i)}_{=T} = T \sum_{i=1}^{d} \omega_i^2. \tag{4.18}$$

Thus, the inequality in (4.18) does not lead to the desired outcome. The uniform boundedness of the first moments of the trace statistic cannot be shown because the upper bound of $E(Z_{T,d})$ is dependent on $T$. Another method must be used to prove the finiteness of the moments of the trace statistic.

[3] The Wishart distribution is a generalization of the gamma distribution. The $W_m(n, \Sigma)$ distribution has a density function when $n \geq m$ (see Muirhead, 1982).

## 4.3 Finiteness of the Moments of the Asymptotic Trace Statistic for $d = 1$

This section demonstrates the finiteness of the moments of the asymptotic trace statistic for $d = K - r = 1$. First, the proof of Lemma 4.1 on the uniform boundedness of the moments of the trace statistic will be presented. Finally, the proof of Theorem 4.1 which shows the uniform integrability of $Z_{T,1}$ and $Z_{T,1}^2$ will complete the proof on the finiteness of the moments of the asymptotic trace statistic.

**Lemma 4.1:** *There are some constants $a$ and $b$ such that, for all $T$,*

(i.) $E(Z_{T,1}^2) < a$,

(ii.) $E(Z_{T,1}^4) < b$.

*Proof.* Assume that $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)'$ is a $T$-dimensional column vector, and $\varepsilon \sim N(0, I_T)$. For $d = 1$ the expression in (4.9) turns into

$$\begin{aligned} Z_{T,1} &= \varepsilon' BA\varepsilon(\varepsilon' A'B'BA\varepsilon)^{-1}\varepsilon' A'B'\varepsilon && (4.19)\\ &= \frac{(\varepsilon' D\varepsilon)^2}{(\varepsilon' D'D\varepsilon)}, && (4.20) \end{aligned}$$

with

$$D = BA = \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 \\ 1 & 0 & \ldots & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \ldots & 1 & 0 \end{pmatrix};$$

$$D'D = \begin{pmatrix} T-1 & T-2 & \ldots & 1 & 0 \\ T-2 & T-2 & \ldots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \ldots & 1 & \vdots \\ 0 & 0 & \ldots & \ldots & 0 \end{pmatrix}.$$

Note that $\varepsilon' D\varepsilon = \varepsilon' D'\varepsilon = \varepsilon' \left(\frac{D+D'}{2}\right)\varepsilon$. As a result (4.20) can be represented by

$$Z_{T,1} = \frac{\left[\varepsilon'(\frac{D+D'}{2})\varepsilon\right]^2}{(\varepsilon' D'D\varepsilon)} = \frac{(\varepsilon' S\varepsilon)^2}{(\varepsilon' F\varepsilon)}, \qquad (4.21)$$

in which $S = \frac{D+D'}{2}$ and $F = D'D$ are $(T \times T)$ symmetric matrices. Moreover, $F$ is a positive semidefinite matrix with $\mathrm{rank}(F) = T-1$, and the eigenvalues $\lambda_1 \geq ... \geq \lambda_{T-1} > \lambda_T = 0$. Furthermore,

$$S = \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & \dots & 0 \end{pmatrix}. \tag{4.22}$$

It is clear from (4.21) that we are dealing with a ratio of quadratic forms in normal variables[4].

To prove (i.) we first apply the Cauchy-Schwarz inequality to $E(Z_{T,1})$.

$$E(Z_{T,1}) \quad = \quad E\left[\frac{(\varepsilon' S\varepsilon)^2}{\varepsilon' F\varepsilon}\right] \leq \sqrt{E\left[(\varepsilon' S\varepsilon)^4\right] E\left[\frac{1}{(\varepsilon' F\varepsilon)^2}\right]} \tag{4.23}$$

To compute $E\left[(\varepsilon' S\varepsilon)^4\right]$, we rewrite (4.22) as

$$S = \frac{1}{2}(J_T - I_T), \tag{4.24}$$

in which $J_T$ is a $(T \times T)$ matrix of ones. The matrix $S$ can be also formulated by

$$\begin{aligned} S \quad &= \quad \frac{1}{2}\left(T\frac{J_T}{T} - \left(I_T - \frac{J_T}{T}\right) - \frac{J_T}{T}\right) && (4.25) \\ &= \quad \frac{1}{2}(TP - Q - P) = \left(\frac{T-1}{2}\right)P - \left(\frac{1}{2}\right)Q, && (4.26) \end{aligned}$$

with $P = \frac{J_T}{T}$ and $Q = (I_T - P)$. The $(T \times T)$ projection matrices $P$ and $Q$ are symmetric and idempotent matrices, i.e. $P' = P$, $PP = P$, $Q = Q'$, $QQ = Q$, and they are orthogonal to each other, i.e. $PQ = 0$. In addition, $\mathrm{rank}(P) = \mathrm{tr}(P) = 1$ and $\mathrm{rank}(Q) = \mathrm{tr}(Q) = T - 1$.

[4]Remark: It is not possible to take the expectation of the two quadratic forms $(\varepsilon' S\varepsilon)^2$ and $1/(\varepsilon' F\varepsilon)$, separately as they are not independent from each other. According to Johnson and Kotz (1970a) for the independence of these two quadratic forms, $SF = 0$ must be fulfilled. In addition to this, to take the expectation of $(\varepsilon' S\varepsilon)^2$ and $(\varepsilon' F\varepsilon)$ separately, the ratio $(\varepsilon' S\varepsilon)^2/(\varepsilon' F\varepsilon)$ must be independent of $(\varepsilon' F\varepsilon)$. This is fulfilled if the matrix $F$ is positive definite, $E(\varepsilon) = 0$ and the eigenvalues of matrix $F$ are all equal (see Jones, 1987; Conniffe and Spencer, 2001). However, we will see later in detail that these conditions are also not achieved.

Equation (4.25) may be seen as the spectral decomposition of the matrix $S$ which reveals that the first eigenvalue of $S$ is $\frac{T-1}{2}$ with multiplicity 1, and the second eigenvalue is $-\frac{1}{2}$ with multiplicity $(T-1)$, i.e.

$$\lambda_1 = \frac{T-1}{2}, \quad \text{and} \quad \lambda_2 = ... = \lambda_T = -\frac{1}{2}. \tag{4.27}$$

This shows us that $S$ is an indefinite matrix. The quadratic form

$$\varepsilon' S \varepsilon \;=\; \left(\frac{T-1}{2}\right)\varepsilon' P \varepsilon - \left(\frac{1}{2}\right)\varepsilon' Q \varepsilon = z_1 - z_2, \tag{4.28}$$

is a combination of two quadratic forms with $z_1 = \left(\frac{T-1}{2}\right)\varepsilon' P \varepsilon$ and $z_2 = \left(\frac{1}{2}\right)\varepsilon' Q \varepsilon$. Moreover, $\varepsilon' P \varepsilon \sim \chi^2_{(1)}$ and $\varepsilon' Q \varepsilon \sim \chi^2_{(T-1)}$ are mutually independently distributed[5] as $PQ = 0$.

Using Equation (4.28) and the expressions for the first four moments of the $\chi^2$ distribution[6], we get

$$\begin{aligned}
E[(\varepsilon' S \varepsilon)^4] &= E[(z_1 - z_2)^4] && (4.30)\\
&= E[z_1^4 - 4z_1^3 z_2 + 6z_1^2 z_2^2 - 4z_1 z_2^3 + z_2^4] \\
&= \left(\frac{T-1}{2}\right)^4 7 \times 5 \times 3 \times 1 - 4\left(\frac{T-1}{2}\right)^4 5 \times 3 \times 1 \\
&\quad +6\left(\frac{1}{2}\right)\left(\frac{T-1}{2}\right)^3 (T+1)3 \times 1 \\
&\quad -4\left(\frac{1}{2}\right)^2\left(\frac{T-1}{2}\right)^2 (T+3)(T-1) \\
&\quad +\left(\frac{1}{2}\right)^4 (T+5)(T+3)(T+1)(T-1) && (4.31)\\
&= c_1 T^4 + o(T^4), && (4.32)
\end{aligned}$$

in which $c_1 \in (0, \infty)$ is a constant.

[5] Johnson and Kotz (1970b) Chapter 29, pp. 177-178, Mathai and Provost (1992) Theorem 5.1.1, p. 196, Judge et al. (1988) pp. 973-974.

[6] The $\nu$th moment of the $\chi^2$ distribution with $n$ degrees of freedom is (Johnson and Kotz, 1970a, Chapter 17, p. 168):

$$\mu_\nu = 2^\nu \frac{\Gamma\left(\nu + \frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}. \tag{4.29}$$

In (4.29), $\Gamma$ denotes the gamma function with $\Gamma(x+1) = x\Gamma(x)$, and thus $\Gamma(n+1) = n!$ for a natural number $n$.

Next, to compute $E\left[(\varepsilon' F \varepsilon)^{-2}\right]$ we use the spectral decomposition of the symmetric matrix $F$,

$$F = V\Lambda V' = \sum_{t=1}^{T} \lambda_t \mathrm{v}_t \mathrm{v}_t', \tag{4.33}$$

in which the diagonal elements of the $(T \times T)$ matrix $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_T)$ are the eigenvalues of the matrix $F$ and $V = (\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_T)$ is the $(T \times T)$ orthogonal matrix of the eigenvectors corresponding to the eigenvalues of the matrix $F$. Note that $\zeta_t = \varepsilon' \mathrm{v}_t \sim N(0,1)$ *i.i.d.* as $\mathrm{v}_t' \mathrm{v}_s = \begin{cases} 1, & t = s \\ 0, & t \neq s \end{cases}$, and $\varepsilon \sim N(0, I_T)$. Then,

$$\varepsilon' F \varepsilon = \sum_{t=1}^{T} \lambda_t \varepsilon' \mathrm{v}_t \mathrm{v}_t' \varepsilon = \sum_{t=1}^{T} \lambda_t \zeta_t^2, \tag{4.34}$$

with $\zeta_t^2 \sim \chi^2_{(1)}$ for $t = 1, \ldots, T$, and the $\zeta_t^2$'s are mutually independently distributed.

As pointed out earlier the positive semidefinite matrix $F$ has $T$ distinct real nonnegative eigenvalues and the smallest eigenvalue is equal to zero, i.e. $\lambda_1 \geq \lambda_2 \geq .... \geq \lambda_{T-1} > \lambda_T = 0$. Thus,

$$\varepsilon' F \varepsilon = \sum_{t=1}^{T} \lambda_t \zeta_t^2 = \sum_{t=1}^{T-1} \lambda_t \zeta_t^2. \tag{4.35}$$

From this representation it becomes obvious that $\varepsilon' F \varepsilon$ is a weighted sum of $\chi^2$ distributed variables. Additionally, we know that $\sum_{t=1}^{T-1} \lambda_t = \mathrm{tr}(F) = \frac{T(T-1)}{2}$. Delete the last row and column of $F$, and consider the following $((T-1) \times (T-1))$ dimensional principle submatrix of $F$.

$$\begin{aligned} \tilde{F} &= \begin{pmatrix} T-1 & T-2 & \ldots & 2 & 1 \\ T-2 & T-2 & \ldots & 2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2 & 2 & \ldots & 2 & 1 \\ 1 & 1 & \ldots & 1 & 1 \end{pmatrix} \\ &= (-1) \begin{pmatrix} -1 & 1 & 0 & \ldots & 0 \\ 1 & -2 & 1 & \ldots & 0 \\ 0 & 1 & -2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & \ddots & -2 \end{pmatrix}^{-1} \end{aligned} \tag{4.36}$$

From (4.36) it is clear that the principle submatrix of $F$ can be represented as the inverse of a tridiagonal Minkowski matrix[7]. Using the Theorem 2 of Yueh (2005) on the eigenvalues of tridiagonal matrices[8] and taking the inverse of the eigenvalues of the tridiagonal matrix achieved from this theorem, the positive real eigenvalues of $F$ can be formulated as[9]

$$\lambda_t = \frac{1}{2 - 2\cos\left(\frac{(2t-1)\pi}{2T-1}\right)} \quad \text{for} \quad t = 1, \ldots, T-1. \tag{4.38}$$

Hence, the largest eigenvalue (or spectral radius[10]) of the matrix $F$ is

$$\lambda_1 = \lambda_{\max} = \frac{1}{2 - 2\cos\left(\frac{\pi}{2T-1}\right)} = c_2 T^2 + o(T^2), \tag{4.39}$$

in which $c_2 \in (0, \infty)$ is a constant. Moreover,

$$\lambda_{T-1} = \frac{1}{2 - 2\cos\left(\frac{(2T-3)\pi}{2T-1}\right)} \to \frac{1}{4} \quad \text{as} \quad T \to \infty. \tag{4.40}$$

To obtain an upper bound for $E\left[(\varepsilon' F \varepsilon)^{-2}\right]$ we use the following series of

[7]see Neumann, 2000; Yueh, 2006.

[8]Yueh (2005) derived the eigenvalues of some tridiagonal matrices using the symbolic calculus method of Cheng (2003).

[9]Remark: Since

$$\varepsilon' F \varepsilon = \begin{pmatrix} \widetilde{\varepsilon}' & \varepsilon_T \end{pmatrix} \begin{pmatrix} \widetilde{F} & \mathbf{0}' \\ \mathbf{0} & 0 \end{pmatrix} \begin{pmatrix} \widetilde{\varepsilon} \\ \varepsilon_T \end{pmatrix} = \widetilde{\varepsilon}' \widetilde{F} \widetilde{\varepsilon}, \tag{4.37}$$

in which $\widetilde{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{T-1})'$ and $\mathbf{0}$ is a $(T-2)$-dimensional row vector of zeros, the positive eigenvalues of $F$ are equal to the eigenvalues of $\widetilde{F}$.

[10]The spectral radius $\rho(A)$ of a square matrix A is defined to be

$$\rho(A) = \max\{|\lambda|; \lambda \in S(A)\},$$

in which $S(A)$ is the set of all the eigenvalues of A. The largest eigenvalue of a nonnegative definite symmetric matrix is equal to its spectral radius, i.e. $\lambda_{\max} = \rho(A)$.

inequalities[11].

$$\varepsilon' F\varepsilon = \sum_{t=1}^{T-1} \lambda_t \zeta_t^2 \geq \sum_{t=1}^{5} \lambda_t \zeta_t^2 \geq \lambda_5 \sum_{t=1}^{5} \zeta_t^2 = \lambda_5 z_4, \tag{4.43}$$

and thus

$$\frac{1}{\varepsilon' F\varepsilon} \leq \frac{1}{\lambda_5 z_4}. \tag{4.44}$$

Note that $z_4 = \sum_{t=1}^{5} \zeta_t^2 \sim \chi^2_{(5)}$ and $\lambda_5$ is the fifth largest eigenvalue of the matrix $F$. For the last equality in (4.43), we use the information that the $\zeta_t^2$'s are mutually independent $\chi^2_{(1)}$ distributed variables.

Finally, on account of (4.23), (4.32) and (4.44) we obtain

$$E(Z_{T,1}) \leq \sqrt{E[(\varepsilon' S\varepsilon)^4]\frac{1}{\lambda_5^2} E\left[\frac{1}{(z_4)^2}\right]} = \sqrt{\frac{1}{3}[c_1 T^4 + o(T^4)]\frac{1}{\lambda_5^2}}. \tag{4.45}$$

$E\left[(z_4)^{-2}\right] = \frac{1}{3}$ is computed by using the expression for the inverse moments of the (central-)$\chi^2_{(n)}$ distributed variables[12].

Thus, to prove $E(Z_{T,1}) < a^*$, for all $T$ with $a^* \in (0, \infty)$, it suffices to show that $\lambda_5 = c_4 T^2 + o(T^2)$, which will lead to $(1/\lambda_5^2) = 1/[c_5 T^4 + o(T^4)]$. In other words, we should prove that

$$\frac{\lambda_5}{T^2} = \frac{1}{2\left[1 - \cos\left(\frac{9\pi}{2T-1}\right)\right] T^2} \to c_4 \text{ as } T \to \infty \text{ for some } c_4 \in (0, \infty). \tag{4.46}$$

or

$$f(T) = T^2 \left[1 - \cos\left(\frac{9\pi}{2T-1}\right)\right] \to c^* \text{ as } T \to \infty \text{ for some } c^* \in (0, \infty). \tag{4.47}$$

[11] The standard way of estimating $\varepsilon' F\varepsilon$ by

$$\varepsilon' F\varepsilon = \widetilde{\varepsilon}' \widetilde{F} \widetilde{\varepsilon} \geq \lambda_{\min}(\widetilde{F})\widetilde{\varepsilon}'\widetilde{\varepsilon}, \tag{4.41}$$

does not work because $\lambda_{\min}(\widetilde{F}) \to \frac{1}{4}$ as $T \to \infty$. So that this estimate yields

$$E\left[\frac{1}{(\varepsilon' F\varepsilon)^2}\right] \leq \frac{1}{\lambda^2_{\min}(\widetilde{F})} E\left[\frac{1}{(z_3)^2}\right] = \frac{1}{16(c_3 T^2 + o(T^2))}, \tag{4.42}$$

in which $z_3 \sim \chi^2_{(T-1)}$, $E[(z_3)^{-2}] = \frac{1}{(T-3)(T-5)}$ and $c_3$ is positive and finite.

[12] From Jones (1986, 1987)

$$E[(\chi^2_{(n)})^{-\nu}] = \prod_{j=1}^{\nu} (n - 2j)^{-1} = [(n-2)(n-4)\dots(n-2\nu)]^{-1},$$

and these moments exist if and only if $n > 2\nu$.

Hence,

$$\lim_{T\to\infty} f(T) = \lim_{T\to\infty} \frac{\left[1-\cos\left(\frac{9\pi}{2T-1}\right)\right]}{\frac{1}{T^2}} = \lim_{T\to\infty}\frac{g(T)}{h(T)} \left(= \frac{0”}{“0}\right), \tag{4.48}$$

as $\lim_{T\to\infty} \cos\left(\frac{9\pi}{2T-1}\right) = 1$, and $\lim_{T\to\infty} \frac{1}{T^2} = 0$. By applying l'Hospital rule we obtain

$$\lim_{T\to\infty}\frac{g(T)}{h(T)} = \lim_{T\to\infty}\frac{g'(T)}{h'(T)} = \lim_{T\to\infty}\frac{-2\sin\left(\frac{9\pi}{2T-1}\right)9\pi(2T-1)^{-2}}{-2T^{-3}} \tag{4.49}$$

$$= \lim_{T\to\infty}\frac{9\pi\sin\left(\frac{9\pi}{2T-1}\right)}{4T^{-1}-4T^{-2}+T^{-3}}\left(= \frac{0”}{“0}\right). \tag{4.50}$$

We apply l'Hospital rule once more to (4.50) and get

$$\lim_{T\to\infty}\frac{g''(T)}{h''(T)} = \lim_{T\to\infty}\frac{-2(9\pi)^2\cos\left(\frac{9\pi}{2T-1}\right)(2T-1)^{-2}}{-(4T^{-2}-8T^{-3}+3T^{-4})} \tag{4.51}$$

$$= \lim_{T\to\infty}\frac{2(9\pi)^2\cos\left(\frac{9\pi}{2T-1}\right)}{(4T^{-2}-8T^{-3}+3T^{-4})(2T-1)^2} \tag{4.52}$$

$$= 2(9\pi)^2\frac{1}{16} > 0, \tag{4.53}$$

which establishes the desired result. Thus,

$$E(Z_{T,1}) \leq \sqrt{E\left[(\varepsilon' S\varepsilon)^4\right]\frac{1}{\lambda_5^2}E\left[\frac{1}{(z_4)^2}\right]} \tag{4.54}$$

$$= \sqrt{\frac{1}{3}\frac{\left[c_1T^4+o(T^4)\right]}{\left[c_5T^4+o(T^4)\right]}} \underset{T\to\infty}{\longrightarrow} \sqrt{\frac{c_1}{3c_5}}. \tag{4.55}$$

This implies that $E(Z_{T,1}) < a^*$ for some $a^*$ and this is valid for all $T$ because of the reasons given in Section 4.2. Consequently, the first moments of $Z_{T,1}$ are uniformly bounded, which completes the proof.

Analogous to this proof we apply again the Cauchy-Schwarz inequality and a similar procedure to prove $E(Z^2_{T,1}) < a$, for all $T$ and $a \in (0,\infty)$.

$$E(Z^2_{T,1}) = E\left[\left(\frac{(\varepsilon' S\varepsilon)^2}{\varepsilon' F\varepsilon}\right)^2\right] \leq \sqrt{E\left[(\varepsilon' S\varepsilon)^8\right]E\left[\frac{1}{(\varepsilon' F\varepsilon)^4}\right]} \tag{4.56}$$

The expression for $E\left[(\varepsilon' S\varepsilon)^8\right]$ can be accomplished by using (4.28) and the first eight moments of the $\chi^2$ distribution with $n$ degrees of freedom, which leads to

$$
\begin{aligned}
E[(\varepsilon' S\varepsilon)^8] &= E[(z_1 - z_2)^8] && (4.57)\\
&= E[z_1^8 - 8z_1^7 z_2 + 28 z_1^6 z_2^2 - 56 z_1^5 z_2^3 + 70 z_1^4 z_2^4 \\
&\quad -56 z_1^3 z_2^5 + 28 z_1^2 z_2^6 - 8 z_1 z_2^7 + z_2^8] && (4.58)\\
&= \left(\frac{T-1}{2}\right)^8 15\times 13\times 11\times 9\times 7\times 5\times 3\times 1\\
&\quad -8\left(\frac{T-1}{2}\right)^8 13\times 11\times 9\times 7\times 5\times 3\times 1\\
&\quad +\frac{28}{2}\left(\frac{T-1}{2}\right)^7 (T+1)11\times 9\times 7\times 5\times 3\times 1\\
&\quad -\frac{56}{2^2}\left(\frac{T-1}{2}\right)^6 (T+1)(T+3)9\times 7\times 5\times 3\times 1\\
&\quad +\frac{70}{2^3}\left(\frac{T-1}{2}\right)^5 (T+1)(T+3)(T+5)7\times 5\times 3\times 1\\
&\quad -\frac{56}{2^4}\left(\frac{T-1}{2}\right)^4 (T+1)(T+3)(T+5)(T+7)5\times 3\times 1\\
&\quad +\frac{28}{2^5}\left(\frac{T-1}{2}\right)^3 (T+1)(T+3)(T+5)(T+7)(T+9)3\times 1\\
&\quad -\frac{8}{2^6}\left(\frac{T-1}{2}\right)^2 (T+1)(T+3)(T+5)(T+7)(T+9)\\
&\quad (T+11) + \frac{1}{2^8}(T-1)(T+1)(T+3)(T+5)(T+7)\\
&\quad (T+9)(T+11)(T+13) && (4.59)\\
&= c_6 T^8 + o(T^8) \quad \text{with} \quad c_6 \in (0,\infty). && (4.60)
\end{aligned}
$$

To obtain $E\left[(\varepsilon' F\varepsilon)^{-4}\right]$, we use the following inequality.

$$
E\left[\frac{1}{(\varepsilon' F\varepsilon)^4}\right] \leq \frac{1}{\lambda_9^4} E\left[\frac{1}{(z_5)^4}\right], \tag{4.61}
$$

with $z_5 \sim \chi^2_{(9)}$. After inserting (4.61) into (4.56), the inequality for $E(Z^2_{T,1})$ turns into

$$
E(Z^2_{T,1}) \leq \sqrt{E\left[(\varepsilon' S\varepsilon)^8\right] \frac{1}{\lambda_9^4} E\left[\frac{1}{(z_5)^4}\right]}. \tag{4.62}
$$

We compute $E\left[(z_5)^{-4}\right]$ as

$$
E\left[(z_5)^{-4}\right] = \frac{1}{(9-2)(9-4)(9-6)(9-8)} = \frac{1}{105}, \tag{4.63}
$$

which is finite and of course independent of $T$. To complete the proof of (i.), we have to demonstrate that $\lambda_9 = c_7 T^2 + o(T^2)$, in other words

$$\frac{\lambda_9}{T^2} = \frac{1}{2\left[1-\cos\left(\frac{17\pi}{2T-1}\right)\right]T^2} \to c_7 \quad \text{as } T \to \infty \text{ and } c_7 \in (0,\infty). \tag{4.64}$$

Therefore, consider

$$\lim_{T\to\infty} f(T) = \lim_{T\to\infty} \frac{\left[1-\cos\left(\frac{17\pi}{2T-1}\right)\right]}{\frac{1}{T^2}} = \lim_{T\to\infty} \frac{g(T)}{h(T)} \left(= \frac{0\text{''}}{\text{``}0}\right). \tag{4.65}$$

Since the limit of $f(T) = g(T)/h(T)$ is $0/0$, the l'Hospital rule is used.

$$\begin{aligned}
\lim_{T\to\infty} \frac{g(T)}{h(T)} &= \lim_{T\to\infty} \frac{g'(T)}{h'(T)} = \lim_{T\to\infty} \frac{-2\sin\left(\frac{17\pi}{2T-1}\right)17\pi(2T-1)^{-2}}{-2T^{-3}} && (4.66)\\
&= \lim_{T\to\infty} \frac{17\pi\sin\left(\frac{17\pi}{2T-1}\right)}{4T^{-1}-4T^{-2}+T^{-3}} \left(= \frac{0\text{''}}{\text{``}0}\right) && (4.67)
\end{aligned}$$

Using the l'Hospital rule once again, we can show that

$$\begin{aligned}
\lim_{T\to\infty} \frac{g''(T)}{h''(T)} &= \lim_{T\to\infty} \frac{-2(17\pi)^2\cos\left(\frac{17\pi}{2T-1}\right)(2T-1)^{-2}}{-(4T^{-2}-8T^{-3}+3T^{-4})} && (4.68)\\
&= \lim_{T\to\infty} \frac{2(17\pi)^2\cos\left(\frac{17\pi}{2T-1}\right)}{(4T^{-2}-8T^{-3}+3T^{-4})(2T-1)^2} && (4.69)\\
&= 2(17\pi)^2\frac{1}{16} > 0. && (4.70)
\end{aligned}$$

As a result

$$\frac{1}{\lambda_9^4} = \frac{1}{c_8 T^8 + o(T^8)} \quad \text{with} \quad c_8 \in (0,\infty). \tag{4.71}$$

This completes the proof of (i.) because

$$\begin{aligned}
E(Z_{T,1}^2) &\leq \sqrt{E\left[(\varepsilon' S\varepsilon)^8\right]\frac{1}{\lambda_9^4}E\left[\frac{1}{(z_5)^4}\right]} && (4.72)\\
&= \sqrt{\frac{1}{105}\frac{\left[c_6T^8+o(T^8)\right]}{\left[c_8T^8+o(T^8)\right]}} \underset{T\to\infty}{\longrightarrow} \sqrt{\frac{c_6}{105c_8}}. && (4.73)
\end{aligned}$$

This implies that $E(Z_{T,1}^2) < a$, for all $T$ with some positive and finite $a$.

The proof of (ii.) is analogous to the proof of (i.), and omitted here to save space.

For the finiteness of the moments of the asymptotic trace statistic we need to prove the following Theorem.

**Theorem 4.1:** *It holds that* $E(Z_1^2) < \infty$, $E(Z_{T,1}) \to E(Z_1)$ and $E(Z_{T,1}^2) \to E(Z_1^2)$ *as* $T \to \infty$.

*Proof.* We know that the trace statistic $Z_{T,1}$ converges in distribution to the asymptotic trace statistic (see Johansen, 1995). Therefore, the second moment of $Z_1$ exists (with $E(Z_{T,1}^2) \to E(Z_1^2)$) if $\{Z_{T,1}^2\}$ is uniformly integrable (see Theorem A on p.14 in Serfling, 1980). A sufficient condition for the uniform integrability of $\{Z_{T,1}^2\}$ is that $E|Z_{T,1}|^{2+\delta}$ is uniformly bounded for some $\delta > 0$, i.e $\sup_T E|Z_{T,1}|^{2+\delta} < \infty$. On account of Lemma 4.1, $E(Z_1^2)$ exists as $E(Z_{T,1}^2) < a < \infty$ and $E(Z_{T,1}^4) < b < \infty$ for all $T$, which completes the proof.

## 4.4 Conclusions

In this chapter we presented the reasons why the proof of Lemma 1 of Larsson et al. (2001) on the existence and finiteness of the moments of the asymptotic trace statistic is not correct.

Note that in Section 4.2 we demonstrated that the upper bound of the first moments of the trace statistic is dependent on $T$. This may be due to the fact that the approximation used for the proof of the general case is inappropriate.

Additionally, the proof of Larsson et al. (2001) was corrected for the case, in which the difference between the number of variables ($K$) and the number of the cointegrating relations ($r$) is one i.e. $d = K - r = 1$. Note that for $d = 1$ the trace statistic reduces to a ratio of quadratic forms.

The proof for the existence and finiteness of the moments of the asymptotic trace statistic for $d = 1$ was based on the Cauchy-Schwarz inequality and a sufficient condition for the uniform integrability of the random variables. Thus, the moments of the trace statistic converge to the moments of the asymptotic trace statistic as the number of time observations goes to infinity.

# Chapter 5

# Testing in the Presence of a Time Trend

Most macroeconomic variables, e.g. prices, gross domestic product, consumption etc., exhibit a trending behavior. To model this behavior in the multivariate time series literature a drift parameter is included in the VAR model. Building on this idea, Saikkonen and Lütkepohl (2000a) proposed LM and LR cointegration tests for data with a linear time trend which are different from the popular Johansen (1995) test. Saikkonen and Lütkepohl (2000a) based their test on the idea of subtracting the estimates of the deterministic terms from the data and applying the cointegration test on the trend-adjusted data. The principle of subtracting the estimates of the deterministic terms of the model was first suggested by Stock and Watson (1988). Following this, Saikkonen and Lütkepohl (2000a) proposed to estimate the deterministic terms using the GLS method. In their simulation study, Saikkonen and Lütkepohl (2000a) concluded that their LR test has better properties than the tests of Johansen (1995). Moreover, by construction under the null hypothesis the limit distribution of their test does not depend on the properties of the deterministic terms.

So far there are only a few examples of maximum-likelihood based panel cointegration tests which allow for a deterministic linear trend in the data generating process. Larsson et al. (2001), who extended the Johansen trace test to panel data and Breitung (2005), who based his tests on the procedure of Saikkonen (1999), showed in their studies that their panel cointegration tests can be extended to the case with deterministic terms, but they did not deliver any proof of their asymptotic results. Additionally, Anderson et al. (2006) introduced a systems panel cointegration test, which allows for a linear time trend. This test is built on the method of Box and Tiao (1977) in which the number of stochastic common trends is determined by the number of eigenvalues close to one. However, there is no maximum-likelihood-based panel cointegration test that relies on the idea of subtracting the estimated deterministic terms prior to testing for cointegration. In order to close this gap we extend the trend-adjusting procedure defined in Saikkonen and Lütkepohl (2000a) to the panel data framework and propose an LR panel cointegration test statistic in the presence of a linear time trend in the DGP. With this new maximum-likelihood-based panel cointegration test statistic one can test for the number of cointegrating relations in the system. This is advantageous compared to the residual-based tests, which can only be used to determine whether there is a cointegrating relation or not.

The chapter is organized as follows: In Section 5.1 the heterogeneous panel data VEC model with linear time trend is introduced. Section 5.2 explains the estimation of the deterministic terms. Section 5.3 presents the new LR panel cointegration test. The size and size-adjusted power properties are examined by means of a Monte Carlo study in Section 5.4. Finally, Section 5.5 gives a summary of the main results. The mathematical Appendix in Section 5.6 contains the proofs on the asymptotic moments.

## 5.1 The Model

For a $K$-dimensional process $y_{it} = (y_{1it}, \ldots, y_{Kit})'$ we consider the following heterogeneous VAR model with linear time trend.

$$\begin{aligned} y_{it} &= \mu_{0i} + \mu_{1i} t + x_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, && (5.1) \\ x_{it} &= A_{i1} x_{i,t-1} + \ldots + A_{i,p_i} x_{i,t-p_i} + \varepsilon_{it}. && (5.2) \end{aligned}$$

Here $\mu_{0i}$ and $\mu_{1i}$ are unknown $K$-dimensional parameter vectors, $p_i$ is the lag order of the VAR process for the $i$th cross-section and $A_{i1}, ..., A_{i,p_i}$ are the $(K \times K)$ unknown coefficient matrices. Moreover, we assume that the $K$-dimensional random errors $\varepsilon_{it}$ are serially and cross-sectionally independent with $\varepsilon_{it} \sim N_K(0, \Omega_i)$, for some nonrandom positive definite matrix $\Omega_i$. For simplicity the initial value condition $x_{it} = 0$, $t \leq 0$ and $i = 1, \ldots, N$, is imposed. However, the results remain valid if we assume that the initial values are drawn from a fixed probability distribution, which does not depend on the sample size.

By subtracting $x_{i,t-1}$ from both sides of (5.2) and rearranging terms we get the VEC form of the process $x_{it}$.

$$\Delta x_{it} = \Pi_i x_{i,t-1} + \sum_{j=1}^{p_i - 1} \Gamma_{ij} \Delta x_{i,t-j} + \varepsilon_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \quad (5.3)$$

in which $\Pi_i = -(I_K - A_{i1} - \ldots - A_{i,p_i})$ and $\Gamma_{ij} = -(A_{i,j+1} + \ldots + A_{i,p_i})$ for $j = 1, \ldots, p_i - 1$. The components of the process $x_{it}$ are assumed to be integrated at most of order one and cointegrated with cointegrating rank $r_i$, $0 \leq r_i \leq K$. In other words, $y_{it}$ is at most $I(1)$ and cointegrated at most of order $r_i$. Thus, the matrix $\Pi_i$ can be decomposed as

$$\Pi_i = \alpha_i \beta_i', \quad i = 1, \ldots, N. \tag{5.4}$$

Note that $\alpha_i$ is the loading and $\beta_i$ is the cointegrating matrix. Both $\alpha_i$ and $\beta_i$ are $(K \times r_i)$ matrices of full column rank.

On account of (5.1), (5.2) and (5.3) we obtain the VEC form of $y_{it}$.

$$\begin{aligned} \Delta y_{it} &= \nu_i + \alpha_i \left[\beta_i' y_{i,t-1} - \tau_i (t-1)\right] + \sum_{j=1}^{p_i - 1} \Gamma_{ij} \Delta y_{i,t-j} + \varepsilon_{it}, \\ & i = 1, ..., N; \quad t = p_i + 1, p_i + 2, ..., T, \end{aligned} \tag{5.5}$$

with $\nu_i = -\Pi_i \mu_{0i} + (I_K - \Gamma_{i1} - \ldots - \Gamma_{i,p_i - 1})\mu_{1i}$ and $\tau_i = \beta_i' \mu_{1i}$.

To determine the number of cointegrating relations among the components of the process $y_{it}$, the rank of the matrix $\Pi_i$ should be tested. The relevant null and alternative hypotheses for the cointegration tests are

$$H_0 : \text{rank}(\Pi_i) = r_i \leq r, \quad i = 1, \ldots, N \tag{5.6}$$

vs.

$$H_1 : \text{rank}(\Pi_i) = K, \quad i = 1, \ldots, N. \tag{5.7}$$

Under the null hypothesis all the cross-sections have at most cointegrating rank $r$, whereas under the alternative hypothesis the rank of $\Pi_i$, $i = 1, \ldots, N$, is $K$. Before testing for the cointegrating rank the data should be trend adjusted. For the trend-adjustment, estimations of the deterministic terms $\mu_{0i}$ and $\mu_{1i}$ are required.

## 5.2 Estimation of the Deterministic Terms

To estimate the parameters $\mu_{0i}$ and $\mu_{1i}$, the GLS method is applied. The data series is then trend-adjusted by subtracting the estimated deterministic terms from $y_{it}$.

For estimating the deterministic terms, we use the initial value condition $x_{it} = 0$, for $t \leq 0$. First we rewrite (5.1) as

$$A_i(L) y_{it} = G_{it} \mu_{0i} + H_{it} \mu_{1i} + \varepsilon_{it}, \quad i = 1, ..., N; \quad t = 1, ..., T, \tag{5.8}$$

with $A_i(L) = I_K - A_{i1}L - ... - A_{i,p_i}L^{p_i}$, $G_{it} = A_i(L)a_t$, $H_{it} = A_i(L)b_t$ and

$$a_t = \begin{cases} 1 & \text{for } t \geq 1 \\ 0 & \text{for } t \leq 0 \end{cases}, \qquad b_t = \begin{cases} t & \text{for } t \geq 1 \\ 0 & \text{for } t \leq 0 \end{cases}.$$

Then, (5.8) is premultiplied by $Q_i'$ which is

$$Q_i = \left[\Omega_i^{-1}\alpha_i(\alpha_i'\Omega_i^{-1}\alpha_i)^{-1/2} : \alpha_{i_\perp}(\alpha_{i_\perp}'\Omega_i\alpha_{i_\perp})^{-1/2}\right] \quad \text{and} \quad Q_iQ_i' = \Omega_i^{-1}, \tag{5.9}$$

so that the resulting error terms $Q_i'\varepsilon_{it}$ have an identity covariance matrix.

Replacing the unknown parameter matrices $\alpha_i$, $\beta_i$, $\Gamma_{ij}$ and $\Omega_i$ of the transformed model by their reduced rank (RR) estimates ($\widetilde{\alpha}_i$, $\widetilde{\beta}_i$, $\widetilde{\Gamma}_{ij}$ and $\widetilde{\Omega}_i$, respectively) from (5.5), the model can be written in a feasible form. Note that the unknown parameters are estimated under the null hypothesis that the cointegrating rank is $r$.

With the estimates of the matrices $\alpha_i$, $\beta_i$, $\Gamma_{ij}$ and their definitions from the previous section, the $(K \times K)$ unknown coefficient matrices $A_{ij}$, $i = 1, \ldots, N$ and $j = 1, .., p_i$, can be estimated by

$$\begin{aligned} \widetilde{A}_{i1} &= I_K + \widetilde{\alpha}_i\widetilde{\beta}_i' + \widetilde{\Gamma}_{i1}, \\ \widetilde{A}_{ij} &= \widetilde{\Gamma}_{ij} - \widetilde{\Gamma}_{i,j-1}, \quad \text{for} \quad j = 2, ..., p_i - 1, \\ \widetilde{A}_{i,p_i} &= -\widetilde{\Gamma}_{i,p_i-1}, \end{aligned}$$

which allows to use the following.

$$\begin{aligned} \widetilde{A}_i(L) &= I_K - \widetilde{A}_{i1}L - ... - \widetilde{A}_{i,p_i}L^{p_i}, \\ \widetilde{G}_{it} &= \widetilde{A}_i(L)a_t \quad \text{and} \\ \widetilde{H}_{it} &= \widetilde{A}_i(L)b_t. \end{aligned}$$

This leads to a feasible form of the transformed model. $\widetilde{\alpha}_{i_\perp}$ and $\widetilde{\beta}_{i_\perp}$ can be obtained from the estimates $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$, respectively. To estimate $Q_i$, the estimates $\widetilde{\alpha}_i$, $\widetilde{\alpha}_{i_\perp}$, $\widetilde{\Omega}_i$ are inserted into (5.9), so that

$$\widetilde{Q}_i = \left[\widetilde{\Omega}_i^{-1}\widetilde{\alpha}_i(\widetilde{\alpha}_i'\widetilde{\Omega}_i^{-1}\widetilde{\alpha}_i)^{-1/2} : \widetilde{\alpha}_{i_\perp}(\widetilde{\alpha}_{i_\perp}'\widetilde{\Omega}_i\widetilde{\alpha}_{i_\perp})^{-1/2}\right] \quad \text{for} \quad i = 1, \ldots, N. \tag{5.10}$$

Finally, the estimators of $\mu_{0i}$ and $\mu_{1i}$ can be obtained by the multivariate least squares method applied to the following auxiliary regression equations, separately for each cross-section.

$$\widetilde{Q}_i'\widetilde{A}_i(L)y_{it} = \widetilde{Q}_i'\widetilde{G}_{it}\mu_{0i} + \widetilde{Q}_i'\widetilde{H}_{it}\mu_{1i} + \widetilde{Q}_i'\varepsilon_{it}, \quad i = 1, ..., N; \quad t = 1, ..., T. \tag{5.11}$$

As pointed out earlier, the least squares estimates of $\widetilde{\mu}_{0i}$ and $\widetilde{\mu}_{1i}$ from (5.11) are used to trend adjust the data, before testing for cointegration.

## 5.3 Panel Cointegration Test

Saikkonen and Lütkepohl (2000a) introduced both LM and LR cointegration test statistics. By means of a simulation study they concluded that the LR tests are preferable to LM tests. Based on this result we propose an LR panel cointegration test statistic, which is an extension of the $\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}}$ statistic of Saikkonen and Lütkepohl (2000a) to panel data.

The new test statistic is based on the following trend-adjusted VECM.

$$\Delta\widetilde{x}_{it} = \Pi_i \widetilde{x}_{i,t-1} + \sum_{j=1}^{p_i-1} \Gamma_{ij}\Delta\widetilde{x}_{i,t-j} + e_{it}, \quad i = 1, ..., N; \quad t = p_i+1, ..., T, \tag{5.12}$$

with $\widetilde{x}_{it} = y_{it} - \widetilde{\mu}_{0i} - \widetilde{\mu}_{1i}t$.

The GLS-based trace statistic for each cross-section is then given by

$$\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}_{iT}}(r) = -2\ln Q_T\{H(r)|H(K)\} = -T\sum_{j=r+1}^{K}\ln(1-\widehat{\lambda}_{ij}). \tag{5.13}$$

Here $\widehat{\lambda}_{i1} \geq \ldots \geq \widehat{\lambda}_{iK}$ are the ordered generalized eigenvalues for cross-section $i$ which are obtained by the eigenvalue problem defined in Johansen (1995).

Under the null hypothesis

$$\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}_{iT}}(r) \xrightarrow{w} \widetilde{Z}_d \quad \text{with} \tag{5.14}$$

$$\widetilde{Z}_d = \mathrm{tr}\left[\left(\int_0^1 W_*(s)\mathrm{d}W_*(s)'\right)'\left(\int_0^1 W_*(s)W_*(s)'\mathrm{d}s\right)^{-1} \left(\int_0^1 W_*(s)\mathrm{d}W_*(s)'\right)\right]. \tag{5.15}$$

$W_*(s) = W(s) - sW(1)$ is a d-dimensional Brownian bridge ($d = K - r$) and $\mathrm{d}W_*(s) = \mathrm{d}W(s) - \mathrm{d}sW(1)$. The proof of the asymptotic distribution of $\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}_{iT}}(r)$ can be found in the Appendix of Saikkonen and Lütkepohl (2000a).

Next, following Larsson et al. (2001), the average of $N$ individual trace statistic is denoted by

$$\overline{\mathrm{LR}}^{\mathrm{GLS}}_{\mathrm{trace}_{NT}}(r) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}_{iT}}(r), \tag{5.16}$$

which is called the $\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}}$-bar statistic. After subtracting the mean and dividing by the standard deviation of the asymptotic trace statistic, the standardized $\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}}$-bar test (henceforth panel SL test) statistic is given by

$$\Upsilon_{\overline{\mathrm{LR}}^{\mathrm{GLS}}_{\mathrm{trace}}} = \frac{\sqrt{N}[\overline{\mathrm{LR}}^{\mathrm{GLS}}_{\mathrm{trace}_{NT}}(r) - E(\widetilde{Z}_d)]}{\sqrt{Var(\widetilde{Z}_d)}}, \tag{5.17}$$

in which $E(\widetilde{Z}_d)$ and $Var(\widetilde{Z}_d)$ are the mean and variance of the individual asymptotic trace statistic in (5.15), respectively. The validity of the test requires that:

**Lemma 5.1:** *The second moment, $E(\widetilde{Z}_d^2)$, exists and is finite.*

The finiteness of the first two moments is necessary to establish the asymptotic distribution of the panel SL test statistic. The proof of Lemma 5.1 for $d = 1$ is presented in Section 5.6 which is analogous to the proof in Section 4.3.

The mean and variance of $\widetilde{Z}_d$ can be approximated by simulation for different values of $d = K - r$ (see Lütkepohl and Saikkonen, 2000). To accomplish this, one generates $T = 1000$ independent $d$-dimensional standard normal variates $\varepsilon_t \sim N(0, I_d)$. Next,

$$A_T = \frac{1}{T^2}\sum_{t=1}^{T}\left[\sum_{m=1}^{t-1}(\varepsilon_m - \bar{\varepsilon})\right]\left[\sum_{m=1}^{t-1}(\varepsilon_m - \bar{\varepsilon})\right]', \tag{5.18}$$

$$B_T = \frac{1}{T}\sum_{t=1}^{T}\left[\sum_{m=1}^{t-1}(\varepsilon_m - \bar{\varepsilon})\right](\varepsilon_t - \bar{\varepsilon})', \tag{5.19}$$

are computed with $\bar{\varepsilon} = T^{-1}\sum_{t=1}^{T}\varepsilon_t$. Since $A_T \xrightarrow{\omega} \int_0^1 W_*(s)W_*(s)'\mathrm{d}s$ and $B_T \xrightarrow{\omega} \int_0^1 W_*(s)\mathrm{d}W_*(s)'$, $\widetilde{Z}_{T,d} = \mathrm{tr}(B_T' A_T^{-1} B_T)$ can serve as an approximation of $\widetilde{Z}_d$ in (5.15). By replicating the experiment 20000 times, the first two moments of the asymptotic $\mathrm{LR}_{\mathrm{trace}}^{\mathrm{GLS}}$ statistic are computed for different $d$ values. The resulting approximations of the mean and variance of $\widetilde{Z}_d$ are represented in Table 5.1.

**Table 5.1:** Simulated first two moments of $\widetilde{Z}_d$.

| $K-r$ | $E(\widetilde{Z}_d)$ | $Var(\widetilde{Z}_d)$ | $K-r$ | $E(\widetilde{Z}_d)$ | $Var(\widetilde{Z}_d)$ |
|---|---|---|---|---|---|
| 1 | 2.69 | 4.38 | 7 | 97.91 | 143.68 |
| 2 | 8.86 | 13.37 | 8 | 127.55 | 187.28 |
| 3 | 18.85 | 28.23 | 9 | 161.20 | 238.00 |
| 4 | 32.78 | 47.94 | 10 | 198.43 | 300.91 |
| 5 | 50.58 | 73.74 | 11 | 239.70 | 357.05 |
| 6 | 72.44 | 105.33 | 12 | 284.87 | 424.86 |

The following theorem is an immediate consequence of the central limit theorem and motivates the procedure.

**Theorem 5.1:** *Under the null hypothesis $H_0$ : rank($\Pi$) $= r_i \leq r$ for all $i = 1, \ldots, N$, the panel cointegration statistic $\Upsilon_{\overline{\mathrm{LR}}^{\mathrm{GLS}}_{\mathrm{trace}}}$ is asymptotically $N(0,1)$ distributed as $T \to \infty$ followed by $N \to \infty$.*

Under certain conditions[1] the asymptotic distribution of the panel cointegration statistic $\Upsilon_{\overline{\mathrm{LR}}^{\mathrm{GLS}}_{\mathrm{trace}}}$ can also be established when $T$ and $N$ tend to infinity jointly.

It is obvious from (5.6)-(5.7) that the panel cointegration test is one-sided, and a test at an asymptotic significance level $\alpha$ rejects $H_0 : \mathrm{rank}(\Pi_i) = r_i \leq r$, for all $i$ if

$$\Upsilon_{\overline{\mathrm{LR}}^{\mathrm{GLS}}_{\mathrm{trace}}}(r) > z_{1-\alpha}.$$

$z_{1-\alpha}$ is the $(1-\alpha)$ quantile of the standard normal distribution.

The sequential testing procedure of Johansen (1988) is implemented to determine the cointegrating rank of the process. First, $H_0 : \mathrm{rank}(\Pi_i) = r_i \leq 0$ is tested. If this null hypothesis is rejected, then $H_0 : \mathrm{rank}(\Pi_i) = r_i \leq 1$ is tested. This procedure continues until the null hypothesis cannot be rejected or $H_0 : \mathrm{rank}(\Pi_i) = r_i \leq K-1$ is rejected. If $H_0 : \mathrm{rank}(\Pi_i) = r_i \leq K-1$ is rejected, then (5.1) is stable.[2]

Following the theory in Larsson (1999) and Larsson et al. (2001) we suggest a second approximation of the moments for the standardization of the panel SL statistic. Larsson (1999) and Larsson et al. (2001) proposed to use the moments from an approximating VAR(1) process, even if the true DGP is a VAR process of higher order. This is motivated by the fact that the moments of the log-likelihood for a VAR(s) process can be approximated sufficiently well by the moments from the log-likelihood for a VAR(1) process, in which $s$ denotes the maximum lag order of the VAR process. In particular, they showed

**Theorem 5.2:** *For all positive integers $n$,*

$$E[(-2\ln Q_T^{(s)})^n] = E[(-2\ln Q_T^{(1)})^n] + O(T^{-1}).$$

*$-2\ln Q_T^{(s)}$ is the maximum log-likelihood for a VAR(s) process and $-2\ln Q_T^{(1)}$ is the maximum log-likelihood for a VAR(1) process, which can be formulated*

[1] see Phillips and Moon, 1999 for the conditions under which the sequential convergence implies joint convergence.

[2] Remark: A VAR($p_i$) process is stable if $\det(A_i(z)) \neq 0$ for $|z| \leq 1$ with $A_i(z) = I_K - A_{i1}z - ... - A_{i,p_i}z^{p_i}$ (see Lütkepohl, 2005).

*as*

$$
\begin{aligned}
-2\ln Q_T^{(1)} &= \operatorname{tr}\left\{ \left[ \sum_{t=1}^{T} \left( \sum_{m=1}^{t-1} (\varepsilon_m - \bar{\varepsilon}) \right) \left( \sum_{m=1}^{t-1} (\varepsilon_m - \bar{\varepsilon}) \right)' \right]^{-1} \right. \\
& \left[ \sum_{t=1}^{T} \left( \sum_{m=1}^{t-1} (\varepsilon_m - \bar{\varepsilon}) \right) (\varepsilon_t - \bar{\varepsilon})' \right] \left[ T^{-1} \sum_{t=1}^{T} (\varepsilon_t - \bar{\varepsilon})(\varepsilon_t - \bar{\varepsilon})' \right]^{-1} \\
& \left. \left[ \sum_{t=1}^{T} (\varepsilon_t - \bar{\varepsilon}) \left( \sum_{m=1}^{t-1} (\varepsilon_m - \bar{\varepsilon}) \right)' \right] \right\} + O_p(T^{-1}),
\end{aligned}
$$

*with* $\varepsilon_t \sim N(0, I_d)$ *and* $\bar{\varepsilon} = T^{-1} \sum_{t=1}^{T} \varepsilon_t$.

Using 50000 replications for different time spans $T$ and values $d = K - r$ the VAR(1) mean and variance are computed by means of the simulation. The results are tabulated in Table 5.2.

**Table 5.2:** Mean and variance values of the VAR(1) approximation.

| $K-r$ | 1 | | 2 | | 3 | | 4 | |
|---|---|---|---|---|---|---|---|---|
| $T-1$ | Mean | $Var$ | Mean | $Var$ | Mean | $Var$ | Mean | $Var$ |
| 10 | 2.11 | 1.75 | 6.60 | 3.50 | 13.21 | 4.69 | 21.65 | 5.27 |
| 25 | 2.42 | 2.95 | 7.77 | 7.42 | 16.01 | 12.63 | 26.98 | 17.82 |
| 50 | 2.53 | 3.54 | 8.28 | 9.90 | 17.34 | 18.31 | 29.61 | 28.41 |
| 100 | 2.61 | 3.90 | 8.59 | 11.44 | 18.15 | 22.70 | 31.27 | 37.21 |
| 200 | 2.66 | 4.21 | 8.76 | 12.49 | 18.56 | 25.27 | 32.10 | 42.87 |
| 500 | 2.67 | 4.21 | 8.86 | 13.25 | 18.85 | 27.17 | 32.57 | 45.76 |
| 1000 | 2.67 | 4.37 | 8.86 | 13.41 | 18.87 | 27.73 | 32.80 | 46.78 |

## 5.4 Monte Carlo Study

Three different DGPs are considered to find out the finite sample properties of the panel SL test. Particular interest is in checking how the test reacts to the changes in the crucial parameters of the three DGPs.

### 5.4.1 DGP A

Since Saikkonen and Lütkepohl (2000a) based their simulation study on the Toda (1994, 1995) process, we consider also a modified version of the Toda process for panel data. The Toda process that we use in this chapter differs slightly from the one in Chapter 3. Here, a linear trend is added to the process and a bivariate system is considered.

For $i = 1, ..., N$ and $t = 1, \ldots, T$, the general form of the Toda process in the presence of a linear trend in the data is

$$y_{it} = \begin{pmatrix} 0 \\ \delta_i \end{pmatrix} + \begin{pmatrix} \psi_a & 0 \\ 0 & \psi_b \end{pmatrix} y_{i,t-1} + \varepsilon_{it},\ \varepsilon_{it} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix}\right) i.i.d. \quad (5.20)$$

Throughout this section the initial values $y_{i0}$ are set to zero. Again, $\theta$ represents the correlation between the innovations to the stationary and nonstationary components of the relevant cross-section. If $\theta \neq 0$, then there is instantaneous correlation between the innovations to the stationary and nonstationary components of the process $y_{it}$. The Toda process is frequently used in the literature because from its canonical form other processes can be obtained by linear transformations of $y_{it}$. This makes the tests under consideration invariant to these transformations.

If $\psi_a = \psi_b = 1$, the true cointegrating rank is zero, and there is no cointegrating relation between the components of the process. Then, (5.20) becomes

$$y_{it} = \delta_i e_2 + I_2 y_{i,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, I_2), \quad (5.21)$$

with $e_2 = \begin{pmatrix} 0 & 1 \end{pmatrix}'$. Thus, the process consists of two nonstationary processes. If $\delta_i \neq 0$, a heterogeneous linear trend parameter is present in the second nonstationary process because in a nonstationary unit root processes a drift parameter generates a linear trend. Moreover, there is no instantaneous correlation between the innovations of the two nonstationary components[3], i.e. $\theta = 0$.

If $|\psi_a| < 1$ and $\psi_b = 1$, the true cointegrating rank of the process is one, and (5.20) can be written as

$$y_{it} = \begin{pmatrix} 0 \\ \delta_i \end{pmatrix} + \begin{pmatrix} \psi_a & 0 \\ 0 & 1 \end{pmatrix} y_{i,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix}\right) i.i.d. \quad (5.22)$$

Hence, the process consists of a stationary and a nonstationary component. Instantaneous correlation is present if $\theta \neq 0$, and in the nonstationary component there is a linear trend for $\delta_i \neq 0$.

If $|\psi_a|, |\psi_b| < 1$, then the true cointegrating rank of the process is two, and the VAR process $y_{it}$ is stable. This can be formulated as

$$y_{it} = \begin{pmatrix} \psi_a & 0 \\ 0 & \psi_b \end{pmatrix} y_{i,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, I_2), \quad (5.23)$$

[3]Since $\theta$ denotes correlation between the innovations to the stationary and nonstationary components of the process, $\theta = 0$.

in which the process consists of two stationary components and $\theta = 0$. $\delta_i \neq 0$ is excluded from the model as a drift parameter will not create a linear time trend for stationary processes. Besides this we obtain the same simulation results even when we include a drift parameter.

Throughout the simulation study, we consider the same values for the parameters $\theta$, $\psi_a$ and $\psi_b$ as in Saikkonen and Lütkepohl (2000a): $\theta \in \{0, 0.8\}$, $\psi_a, \psi_b \in \{0.5, 0.7, 0.8, 0.9, 0.95, 1\}$. The time and cross-section dimensions are the values, which are also taken by Larsson et al. (2001): $N \in \{1, 5, 10, 25, 50\}$ and $T - p \in \{10, 25, 50, 100, 200, 500, 1000\}$, in which $p$ is the VAR order of the underlying DGP[4]. The trend parameter is independently generated from a uniform distribution $\delta_i \sim U(0, 2)$. In addition to this, we also consider that the trend parameter is homogeneous, i.e. $\delta_i = 1$ for all $i$. However, this has no affect on the properties of the test. Indeed, the same results are achieved for both heterogeneous and homogeneous cases.

### 5.4.2 DGP B

The second DGP is a VAR(2) process, which allows for a better examination of the properties of the test based on the VAR(1) approximation of the moments. In particular we see how the test behaves when the underlying VAR process has a higher order than one.

If the true cointegrating rank is zero, the DGP takes on the form

$$y_{it} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0.8 & 0 \\ 0 & 0.4 \end{pmatrix} y_{i,t-1} + \begin{pmatrix} 0.2 & 0 \\ 0 & 0.6 \end{pmatrix} y_{i,t-2} + \varepsilon_{it}, \; \varepsilon_{it} \sim N(0, I_2), \quad (5.24)$$

with

$$\Pi_i = \Pi = -\left( I_2 - \begin{pmatrix} 0.8 & 0 \\ 0 & 0.4 \end{pmatrix} - \begin{pmatrix} 0.2 & 0 \\ 0 & 0.6 \end{pmatrix} \right) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad i = 1, \ldots, N.$$

If the true cointegrating rank is one, the DGP is

$$y_{it} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \psi & 0 \\ 0 & 0.4 \end{pmatrix} y_{i,t-1} + \begin{pmatrix} 0.2 & 0 \\ 0 & 0.6 \end{pmatrix} y_{i,t-2} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, I_2), \quad (5.25)$$

with

$$\Pi_i = \Pi = -\left( I_2 - \begin{pmatrix} \psi & 0 \\ 0 & 0.4 \end{pmatrix} - \begin{pmatrix} 0.2 & 0 \\ 0 & 0.6 \end{pmatrix} \right) = \begin{pmatrix} \psi - 0.8 & 0 \\ 0 & 0 \end{pmatrix}, \; i = 1, \ldots, N.$$

[4]In our study additionally, we consider $T - p \in \{500, 1000\}$ to find out the properties of the tests when T is large.

If $\psi < 0.8$, then the DGP process consists of a stationary and a nonstationary component. To generate the same $\Pi_i$ matrices as in DGP A, the $\psi$ parameter takes on the values $\psi \in \{0.5, 0.6, 0.7, 0.75\}$. The trend parameter takes on the value 1 for all $i$ because a cross-section varying trend term does not affect the results of the simulation study.

A VAR(2) process with a true cointegrating rank of two can be generated as follows.

$$y_{it} = \begin{pmatrix} \psi & 0 \\ 0 & 0.3 \end{pmatrix} y_{i,t-1} + \begin{pmatrix} 0.2 & 0 \\ 0 & 0.2 \end{pmatrix} y_{i,t-2} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, I_2) \tag{5.26}$$

If we assume again that $\psi \in \{0.5, 0.6, 0.7, 0.75\}$, the DGP is composed of two stationary processes. The trend parameter is not included in the expression as this will not generate a linear trend.

### 5.4.3 DGP C

The third DGP considered in this simulation study is that of Breitung (2005). DGP C differs from the other two DGPs in so far as both, the trend parameter and the parameters of the coefficient matrix, are heterogeneous over the cross-section dimension. This is quite suitable for the heterogeneous structure of the model introduced in (5.1) and (5.2). The DGP is based on the following VAR(1) model.

$$y_{it} = \mu_i \left( \begin{array}{c} 1 \\ 1 \end{array} \right) + \left( \begin{array}{cc} 1 - a_{1i} & -a_{1i} b \\ -a_{2i} & 1 - a_{2i} b \end{array} \right) y_{i,t-1} + \varepsilon_{it}, \tag{5.27}$$

in which $\varepsilon_{it} = u_{it} + \vartheta_i u_{i,t-1}$, $u_{it} \sim N(0, I_2)$ $i.i.d$ and $y_{i0} = 0$, $i = 1, \dots, N$. If $\vartheta_i \neq 0$, then there is correlation between the components of the process $y_{it}$. Furthermore, the cross-section varying parameters are generated from uniform distributions: $\mu_i \sim U(0,1)$, $\vartheta_i \sim U(0, 0.5)$, $a_{Ki} \sim U(0.1, 0.5)$ for $K = 1, 2$ and $b = 1$.

### 5.4.4 Simulation Results

In this section the simulation results based on the three different data generating processes explained above are presented. Throughout the simulation study the test statistics are computed with two different approximations, i.e approximation based on asymptotic moments and on VAR(1) moments. Similar to the Monte Carlo study of Breitung (2005), we compare our panel SL test with the panel test of Larsson et al. (2001) allowing a linear time trend (henceforth LLL test), which is an extension of the test of Johansen (1995) with deterministic terms. The LLL panel test statistic is computed using the asymptotic moments presented in Breitung (2005). To include the results for the LLL test based on VAR(1) moments, the moments are calculated analogous to the procedure described in Section 5.3.

Note that the total number of replications is 1000. While generating the random error terms, seeded values are used and the first 50 observations are deleted, so that the starting values are not zero anymore. The tests are programmed in GAUSS 6.0.

### Simulation Results for DGP A

With the approximation based on asymptotic moments the size[5] of the panel SL test for the true cointegrating rank of zero (see Table 5.3) fluctuates between 0.053 (for $T = 25$, $N = 25$) and 0.118 (for $T = 10$, $N = 50$). If the test statistic is approximated with VAR(1) moments, the empirical size of the test is around the 5% level for $T = 500, 1000$ and otherwise it is severely oversized. Even worse, the LLL test is severely oversized for short time dimensions under both approximations, and the distortion increases with the increase in $N$. Moreover, its size comes close to the nominal 5% rejection level for larger time dimensions; it reaches 0.055 for $T = 1000$, $N = 10$, when the VAR(1) moments are used. Overall, based on the asymptotic moments the panel SL test shows the best size properties if the true cointegrating rank of the process is zero. Note that with an increase in $T$, the size results with different approximations converge to each other.

To save space, just the extreme cases, i.e $\psi_a = 0.7$, $\psi_a = 0.95$, are shown for the true cointegrating rank of one. When the asymptotic moments are used to approximate the panel statistics, the true hypothesis of $r = 1$ for $\psi_a = 0.7$ cannot be rejected if $T = 10, 25$ and $N \geq 10$ (see Table 5.4). With the increase in $T$ the size of the panel SL test rises and is around the 5% level for $T \geq 100$, and it varies between 0.056 (for $T = 1000$, $N = 50$) and 0.083 (for $T = 500$, $N = 50$). If VAR(1) moments are used, the size of the panel SL test comes close to the 5% level, for $T = 25$. Moreover, based on the VAR(1) moments the LLL test shows poor size properties for small $T$. However, if $T = 1000$, the size of the LLL test under both approximations is around the 5% level. For $\psi_a = 0.95$, the panel SL test is undersized for almost all $T$ and $N$ combinations, except for $T = 1000$, $N \leq 10$ (see Table 5.4). In the latter case the size is exactly 5% with both approximations. The LLL test is also

[5] In the tables presenting the empirical size results the columns denoted by "asym" refer to the results of the tests based on the moments of the asymptotic trace statistic, whereas the columns denoted by "VAR(1)" present the results of the tests based on VAR(1) moments.

undersized for almost all cases, but the most important difference between the properties of the two tests is that if VAR(1) moments are used, the LLL test is oversized for $T = 10$. With an increase in $T$, the size of the LLL test moves close to the 5% nominal level. However, the panel SL test has once more better size properties than the LLL test when $T$ increases.

As it is apparent from Table 5.5, with the introduction of the correlation between the stationary and nonstationary components of the DGP, when $\psi_a = 0.7$ the panel SL test has reasonable size either for $N \leq 10$ or $T = 1000$, independent of the chosen approximation. Hence, for almost all $T$ and $N$ combinations the size of the panel SL test is zero if the true cointegrating rank is one, $\psi_a = 0.95$ and $N \geq 10$. (see Table 5.5). If the panel SL test statistic is approximated with VAR(1) moments, the test has just the correct size for $T = 25$, $N = 10, 25$ and $T = 1000$, $N = 5$ as $\psi_a = 0.7$. Otherwise the test is size distorted for both $\psi_a$ being either 0.7 or 0.95. However, with the approximation based on asymptotic moments the LLL test is undersized for small $T$. With the increase in $T$ the size approaches the nominal level, and the test becomes oversized with a further increase in $T$ and $N$. The LLL test approximated with VAR(1) moments is again severely oversized for short time periods, and the size moves around the 5% level, but does not approach it even for large $T$. In general, none of the tests have nice size properties for $\psi_a = 0.95$.

In line with Banerjee et al. (2004) we observe nonmonotonicities in the results on the size properties of the tests. The sizes of the tests do not increase or decrease monotonically with the increase in $T$ and/or $N$.

Figures 5.1-5.3 present the size-adjusted power results for DGP A[6]. For the true cointegrating rank of one with $\theta = 0$, it is obvious from Figure 5.1 that the size-adjusted power of the LLL test is slightly better than the size-adjusted power of the panel SL test. For both tests the approximation based on VAR(1) moments lead to higher power than the approximation based on asymptotic moments, and their powers approach unity even for small $T$ if $N$ increases. Just the power properties for small $T$ are presented because the power is almost always unity if $T$ and $N$ are sufficiently large. The same conclusions are also visible in Figure 5.2, in which the true cointegrating rank is one and $\theta = 0.8$.

From Figure 5.3 it can be concluded that if both test statistics are approximated with asymptotic moments, the false hypothesis of one cointegrating relation cannot be rejected for $T = 10$. On the contrary, if the test statistics are approximated with VAR(1) moments the powers of the tests increase

[6]The size-adjusted power results for the true cointegrating rank of zero are not illustrated as the power of the tests for the false hypothesis of one cointegrating relation is always zero.

with an increase in $N$, and the power of the LLL test is higher. For both tests the power is higher if VAR(1) moments are used for approximation. If $\psi_a$ parameter[7] increases, higher $T$ and $N$ are necessary so that the powers of the tests tend to unity. Moreover, the LLL test is the least powerful test for $T = 50$ and $\psi_a = 0.95$.

Please note that the size and size-adjusted power results remain the same if a cross-section invariant trend parameter is assumed, i.e. $\delta_i = 1$ for $i = 1, ..., N$, instead of a heterogeneous one. This outcome coincides with the simulation results of Saikkonen and Lütkepohl (2000a).

In general, for DGP A, the panel SL test has better size properties in comparison to the LLL test under both approximations. On the contrary, the power of the LLL test is the highest when the test statistic is approximated with VAR(1) moments.

### Simulation Results for DGP B

Table 5.6 demonstrates that the panel SL test is oversized for $T \leq 50$ and its size increases with an increase in $N$. For $T \geq 100$ the size of the panel SL test ranges from 0.057 (for $T = 500$, $N = 10$) to 0.093 (for $T = 100$, $N = 50$). But if the test statistic is approximated with VAR(1) moments, the test is oversized for $T \leq 200$, and the size is around the 5% nominal significance level only for $T \geq 500$. The LLL test is always more distorted than the panel SL test independent of the chosen approximation. Moreover, if asymptotic moments are used, the size of the panel SL test approaches the 5% level for $T \geq 100$ and $N < 10$. For the true cointegrating rank of zero the panel SL test has the most reasonable size among the two tests and approximations.

To compare the size of the panel SL and LLL tests for the true cointegrating rank of one, just the results related to the two cases $\psi = 0.5$ and $\psi = 0.75$ are presented because the results for $\psi = 0.6$ and $\psi = 0.7$ lie in between these two cases. In Table 5.7 both tests exhibit similar behavior with the approximation based on asymptotic moments, i.e. they are both undersized for small $T$ and slightly oversized for large $T$. The size of the the LLL test is precisely 0.050 if $T = 100$ and $N = 1$ (no panel data). If the test statistics are approximated with VAR(1) moments, the properties of the tests are different for small $T$. The panel SL test is undersized for $T = 10$, whereas the LLL test is badly oversized. If $T \geq 100$, the size of the panel SL test ranges from 0.050 (for $T = 200$, $N = 5$) to 0.137 (for $T = 100$, $N = 50$), whereas if $T \geq 25$ the size of the LLL test lies in between 0.010 (for $T = 50$, $N = 50$) and 0.106 (for $T = 200, N = 50$).

[7] For DGP A to achieve the true cointegrating rank of two, $\psi_b$ parameter is held constant at 0.5.

**Table 5.3:** Empirical size results of the tests for DGP A and true cointegrating rank of zero.

| | | panel SL | | LLL | |
|---|---|---|---|---|---|
| T-1 | N | asym | VAR(1) | asym | VAR(1) |
| 10 | 1 | 0.072 | 0.372 | 0.188 | 0.765 |
| | 5 | 0.089 | 0.720 | 0.394 | 0.996 |
| | 10 | 0.088 | 0.890 | 0.604 | 1.000 |
| | 25 | 0.104 | 0.999 | 0.906 | 1.000 |
| | 50 | 0.118 | 1.000 | 0.990 | 1.000 |
| 25 | 1 | 0.067 | 0.160 | 0.083 | 0.286 |
| | 5 | 0.073 | 0.286 | 0.114 | 0.611 |
| | 10 | 0.058 | 0.395 | 0.129 | 0.774 |
| | 25 | 0.053 | 0.636 | 0.266 | 0.983 |
| | 50 | 0.075 | 0.836 | 0.379 | 1.000 |
| 50 | 1 | 0.081 | 0.128 | 0.079 | 0.165 |
| | 5 | 0.062 | 0.154 | 0.100 | 0.253 |
| | 10 | 0.067 | 0.181 | 0.100 | 0.376 |
| | 25 | 0.061 | 0.281 | 0.142 | 0.604 |
| | 50 | 0.057 | 0.426 | 0.178 | 0.819 |
| 100 | 1 | 0.064 | 0.076 | 0.064 | 0.092 |
| | 5 | 0.056 | 0.096 | 0.058 | 0.112 |
| | 10 | 0.060 | 0.114 | 0.076 | 0.168 |
| | 25 | 0.077 | 0.160 | 0.119 | 0.284 |
| | 50 | 0.075 | 0.220 | 0.147 | 0.387 |
| 200 | 1 | 0.076 | 0.084 | 0.070 | 0.082 |
| | 5 | 0.061 | 0.071 | 0.070 | 0.080 |
| | 10 | 0.056 | 0.071 | 0.074 | 0.088 |
| | 25 | 0.077 | 0.109 | 0.110 | 0.139 |
| | 50 | 0.074 | 0.105 | 0.124 | 0.155 |
| 500 | 1 | 0.069 | 0.069 | 0.064 | 0.062 |
| | 5 | 0.076 | 0.077 | 0.082 | 0.077 |
| | 10 | 0.074 | 0.074 | 0.079 | 0.072 |
| | 25 | 0.068 | 0.069 | 0.088 | 0.078 |
| | 50 | 0.074 | 0.076 | 0.115 | 0.094 |
| 1000 | 1 | 0.061 | 0.061 | 0.080 | 0.076 |
| | 5 | 0.070 | 0.070 | 0.068 | 0.058 |
| | 10 | 0.066 | 0.066 | 0.066 | 0.055 |
| | 25 | 0.068 | 0.068 | 0.081 | 0.059 |
| | 50 | 0.069 | 0.069 | 0.118 | 0.073 |

**Table 5.4:** Empirical size results of the tests for DGP A and true cointegrating rank of one with $\theta = 0$.

| | | $\psi_a = 0.7$ | | | | $\psi_a = 0.95$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | panel SL | | LLL | | panel SL | | LLL | |
| T-1 | N | asym | VAR(1) | asym | VAR(1) | asym | VAR(1) | asym | VAR(1) |
| 10 | 1 | 0.022 | 0.087 | 0.025 | 0.187 | 0.018 | 0.083 | 0.021 | 0.162 |
| | 5 | 0.008 | 0.095 | 0.002 | 0.324 | 0.003 | 0.086 | 0.004 | 0.311 |
| | 10 | 0.001 | 0.122 | 0.002 | 0.411 | 0.001 | 0.102 | 0.001 | 0.406 |
| | 25 | 0.000 | 0.174 | 0.001 | 0.654 | 0.000 | 0.110 | 0.000 | 0.667 |
| | 50 | 0.000 | 0.238 | 0.000 | 0.858 | 0.000 | 0.102 | 0.000 | 0.860 |
| 25 | 1 | 0.039 | 0.068 | 0.012 | 0.049 | 0.015 | 0.021 | 0.007 | 0.036 |
| | 5 | 0.016 | 0.054 | 0.003 | 0.025 | 0.003 | 0.010 | 0.001 | 0.011 |
| | 10 | 0.006 | 0.042 | 0.001 | 0.020 | 0.000 | 0.002 | 0.000 | 0.006 |
| | 25 | 0.004 | 0.049 | 0.000 | 0.024 | 0.000 | 0.000 | 0.000 | 0.002 |
| | 50 | 0.001 | 0.054 | 0.000 | 0.010 | 0.000 | 0.000 | 0.000 | 0.000 |
| 50 | 1 | 0.062 | 0.080 | 0.030 | 0.063 | 0.015 | 0.024 | 0.000 | 0.019 |
| | 5 | 0.058 | 0.085 | 0.016 | 0.046 | 0.002 | 0.005 | 0.000 | 0.000 |
| | 10 | 0.038 | 0.089 | 0.009 | 0.046 | 0.000 | 0.001 | 0.000 | 0.002 |
| | 25 | 0.038 | 0.106 | 0.002 | 0.042 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 50 | 0.031 | 0.122 | 0.002 | 0.028 | 0.000 | 0.000 | 0.000 | 0.000 |
| 100 | 1 | 0.060 | 0.071 | 0.053 | 0.064 | 0.013 | 0.018 | 0.013 | 0.013 |
| | 5 | 0.069 | 0.092 | 0.036 | 0.063 | 0.001 | 0.003 | 0.000 | 0.000 |
| | 10 | 0.064 | 0.088 | 0.059 | 0.078 | 0.000 | 0.000 | 0.002 | 0.003 |
| | 25 | 0.063 | 0.111 | 0.058 | 0.099 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 50 | 0.079 | 0.149 | 0.048 | 0.119 | 0.000 | 0.000 | 0.000 | 0.000 |
| 200 | 1 | 0.063 | 0.062 | 0.068 | 0.077 | 0.022 | 0.020 | 0.014 | 0.016 |
| | 5 | 0.067 | 0.069 | 0.071 | 0.082 | 0.012 | 0.011 | 0.003 | 0.003 |
| | 10 | 0.059 | 0.068 | 0.069 | 0.081 | 0.003 | 0.003 | 0.003 | 0.004 |
| | 25 | 0.060 | 0.069 | 0.073 | 0.088 | 0.000 | 0.000 | 0.001 | 0.001 |
| | 50 | 0.068 | 0.084 | 0.082 | 0.109 | 0.000 | 0.000 | 0.000 | 0.000 |
| 500 | 1 | 0.055 | 0.057 | 0.063 | 0.064 | 0.041 | 0.044 | 0.049 | 0.049 |
| | 5 | 0.066 | 0.073 | 0.076 | 0.076 | 0.041 | 0.046 | 0.046 | 0.046 |
| | 10 | 0.077 | 0.084 | 0.070 | 0.070 | 0.027 | 0.029 | 0.037 | 0.037 |
| | 25 | 0.064 | 0.072 | 0.068 | 0.067 | 0.012 | 0.014 | 0.025 | 0.025 |
| | 50 | 0.083 | 0.095 | 0.074 | 0.073 | 0.013 | 0.016 | 0.020 | 0.020 |
| 1000 | 1 | 0.065 | 0.065 | 0.066 | 0.066 | 0.050 | 0.050 | 0.067 | 0.067 |
| | 5 | 0.066 | 0.067 | 0.073 | 0.071 | 0.053 | 0.053 | 0.073 | 0.072 |
| | 10 | 0.064 | 0.066 | 0.051 | 0.050 | 0.050 | 0.051 | 0.058 | 0.058 |
| | 25 | 0.069 | 0.072 | 0.071 | 0.068 | 0.034 | 0.035 | 0.073 | 0.071 |
| | 50 | 0.056 | 0.060 | 0.075 | 0.071 | 0.020 | 0.024 | 0.086 | 0.078 |

**Table 5.5:** Empirical size results of the tests for DGP A and true cointegrating rank of one with $\theta = 0.8$.

| | | $\psi_a = 0.7$ | | | | $\psi_a = 0.95$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | panel SL | | LLL | | panel SL | | LLL | |
| T-1 | N | asym | VAR(1) | asym | VAR(1) | asym | VAR(1) | asym | VAR(1) |
| 10 | 1 | 0.029 | 0.113 | 0.026 | 0.236 | 0.019 | 0.081 | 0.022 | 0.162 |
| | 5 | 0.016 | 0.155 | 0.016 | 0.482 | 0.005 | 0.085 | 0.005 | 0.323 |
| | 10 | 0.012 | 0.188 | 0.015 | 0.688 | 0.001 | 0.097 | 0.001 | 0.416 |
| | 25 | 0.003 | 0.294 | 0.008 | 0.911 | 0.000 | 0.100 | 0.000 | 0.679 |
| | 50 | 0.001 | 0.423 | 0.003 | 0.991 | 0.000 | 0.107 | 0.000 | 0.874 |
| 25 | 1 | 0.045 | 0.073 | 0.043 | 0.122 | 0.015 | 0.032 | 0.008 | 0.037 |
| | 5 | 0.019 | 0.068 | 0.034 | 0.190 | 0.003 | 0.007 | 0.000 | 0.013 |
| | 10 | 0.011 | 0.050 | 0.040 | 0.253 | 0.000 | 0.006 | 0.000 | 0.008 |
| | 25 | 0.008 | 0.051 | 0.045 | 0.479 | 0.000 | 0.000 | 0.000 | 0.003 |
| | 50 | 0.003 | 0.044 | 0.059 | 0.710 | 0.000 | 0.000 | 0.000 | 0.000 |
| 50 | 1 | 0.059 | 0.070 | 0.079 | 0.118 | 0.014 | 0.017 | 0.011 | 0.019 |
| | 5 | 0.031 | 0.044 | 0.080 | 0.158 | 0.004 | 0.006 | 0.002 | 0.003 |
| | 10 | 0.019 | 0.032 | 0.086 | 0.215 | 0.000 | 0.000 | 0.001 | 0.005 |
| | 25 | 0.007 | 0.020 | 0.137 | 0.380 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 50 | 0.004 | 0.013 | 0.209 | 0.561 | 0.000 | 0.000 | 0.000 | 0.000 |
| 100 | 1 | 0.041 | 0.055 | 0.074 | 0.094 | 0.016 | 0.017 | 0.023 | 0.034 |
| | 5 | 0.033 | 0.042 | 0.093 | 0.127 | 0.000 | 0.001 | 0.010 | 0.017 |
| | 10 | 0.015 | 0.027 | 0.079 | 0.125 | 0.000 | 0.000 | 0.010 | 0.013 |
| | 25 | 0.008 | 0.018 | 0.116 | 0.198 | 0.000 | 0.000 | 0.004 | 0.011 |
| | 50 | 0.005 | 0.015 | 0.176 | 0.287 | 0.000 | 0.000 | 0.000 | 0.003 |
| 200 | 1 | 0.037 | 0.040 | 0.079 | 0.088 | 0.013 | 0.017 | 0.049 | 0.054 |
| | 5 | 0.039 | 0.044 | 0.077 | 0.094 | 0.003 | 0.000 | 0.056 | 0.069 |
| | 10 | 0.027 | 0.032 | 0.067 | 0.082 | 0.000 | 0.000 | 0.047 | 0.058 |
| | 25 | 0.017 | 0.021 | 0.091 | 0.122 | 0.000 | 0.000 | 0.078 | 0.101 |
| | 50 | 0.008 | 0.008 | 0.120 | 0.161 | 0.000 | 0.000 | 0.099 | 0.150 |
| 500 | 1 | 0.071 | 0.059 | 0.068 | 0.069 | 0.017 | 0.018 | 0.080 | 0.084 |
| | 5 | 0.054 | 0.053 | 0.068 | 0.069 | 0.003 | 0.000 | 0.092 | 0.093 |
| | 10 | 0.039 | 0.056 | 0.083 | 0.084 | 0.001 | 0.000 | 0.109 | 0.109 |
| | 25 | 0.032 | 0.029 | 0.082 | 0.081 | 0.000 | 0.000 | 0.137 | 0.137 |
| | 50 | 0.019 | 0.035 | 0.096 | 0.092 | 0.000 | 0.000 | 0.178 | 0.175 |
| 1000 | 1 | 0.060 | 0.061 | 0.071 | 0.071 | 0.019 | 0.015 | 0.080 | 0.080 |
| | 5 | 0.049 | 0.049 | 0.062 | 0.061 | 0.002 | 0.002 | 0.070 | 0.069 |
| | 10 | 0.045 | 0.047 | 0.065 | 0.062 | 0.000 | 0.001 | 0.086 | 0.084 |
| | 25 | 0.044 | 0.045 | 0.061 | 0.059 | 0.000 | 0.000 | 0.081 | 0.077 |
| | 50 | 0.036 | 0.038 | 0.082 | 0.079 | 0.000 | 0.000 | 0.148 | 0.137 |

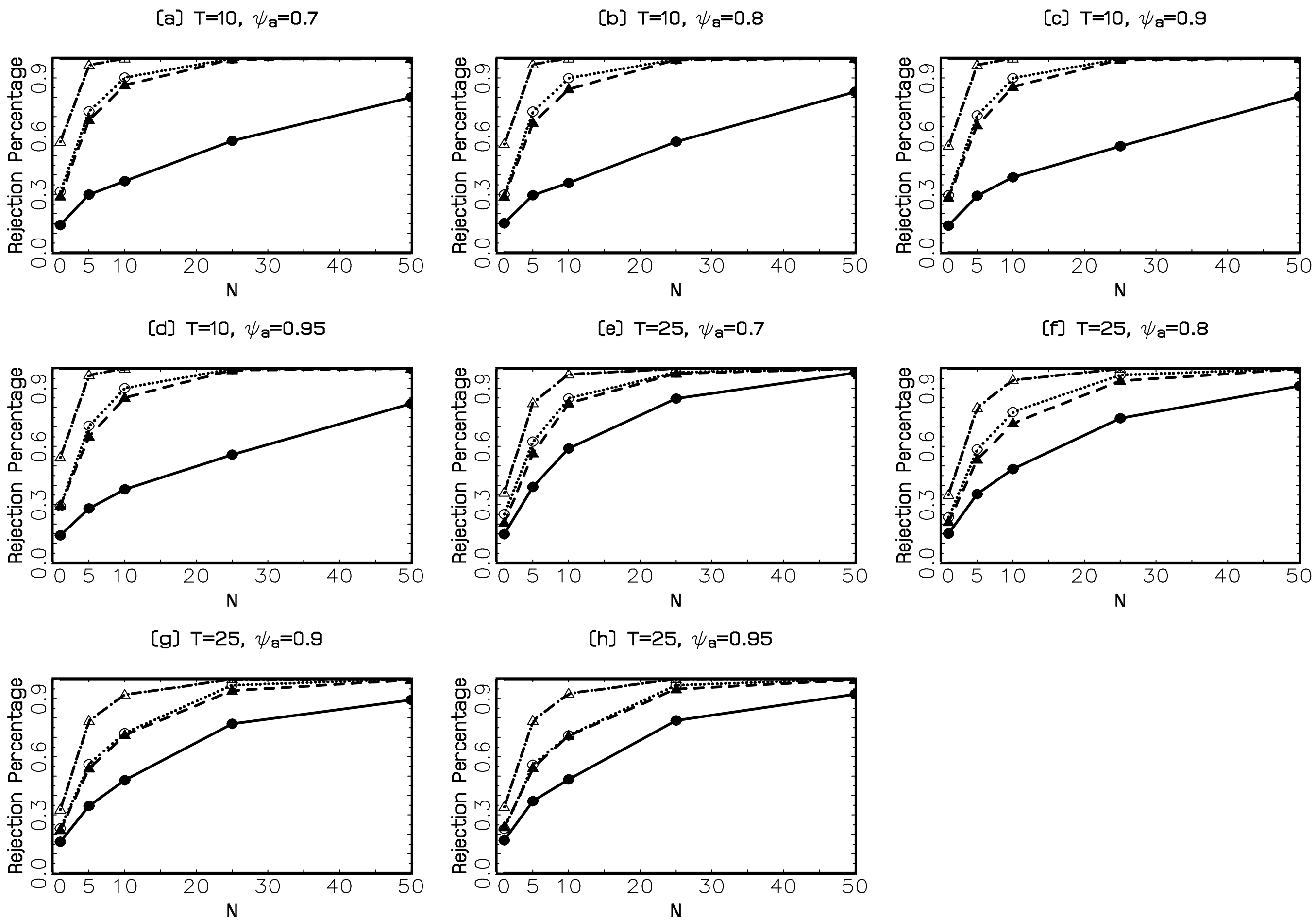


**Figure 5.1:** Small sample size-adjusted power results of the tests for DGP A and true cointegrating rank of one with $\theta = 0$. •——— panel SL-asym, ▲ – – – – panel SL-VAR(1), ∘........ LLL-asym, △ – · – · –· LLL-VAR(1).

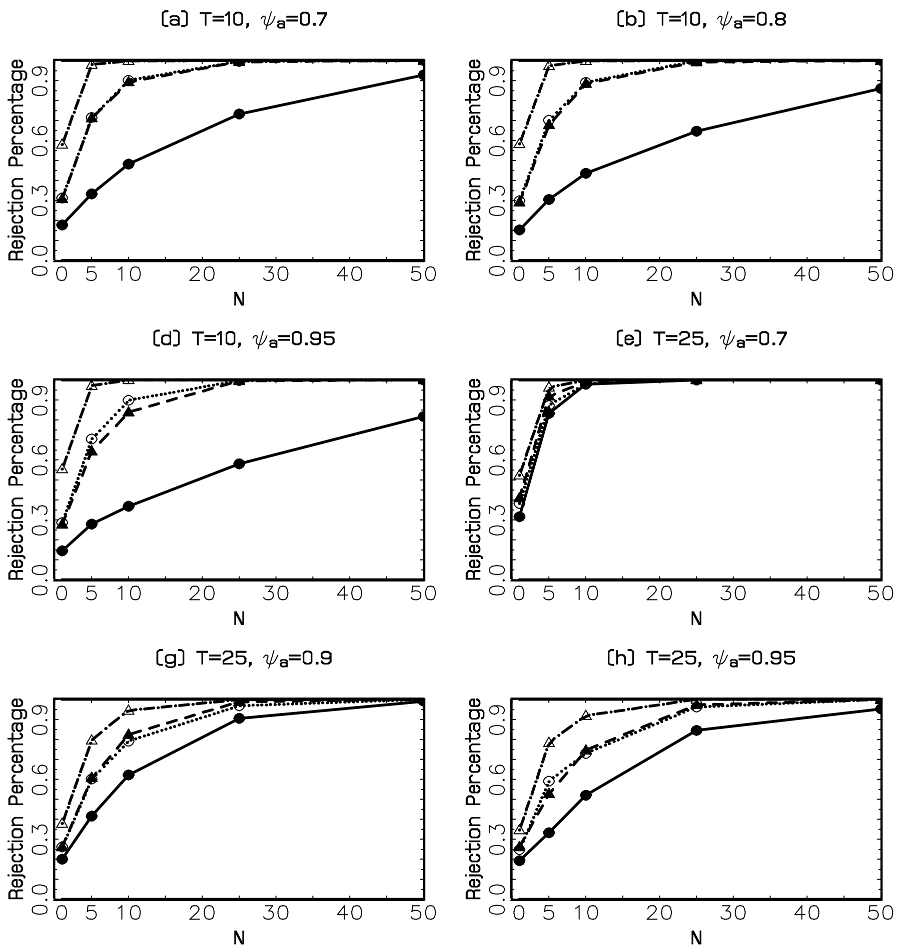


**Figure 5.2:** Small sample size-adjusted power results of the tests for DGP A and true cointegrating rank of one with $\theta = 0.8$. $\bullet$——— panel SL-asym, $\blacktriangle$ − − − − panel SL-VAR(1), $\circ$........ LLL-asym, $\triangle$ − · − · −· LLL-VAR(1).

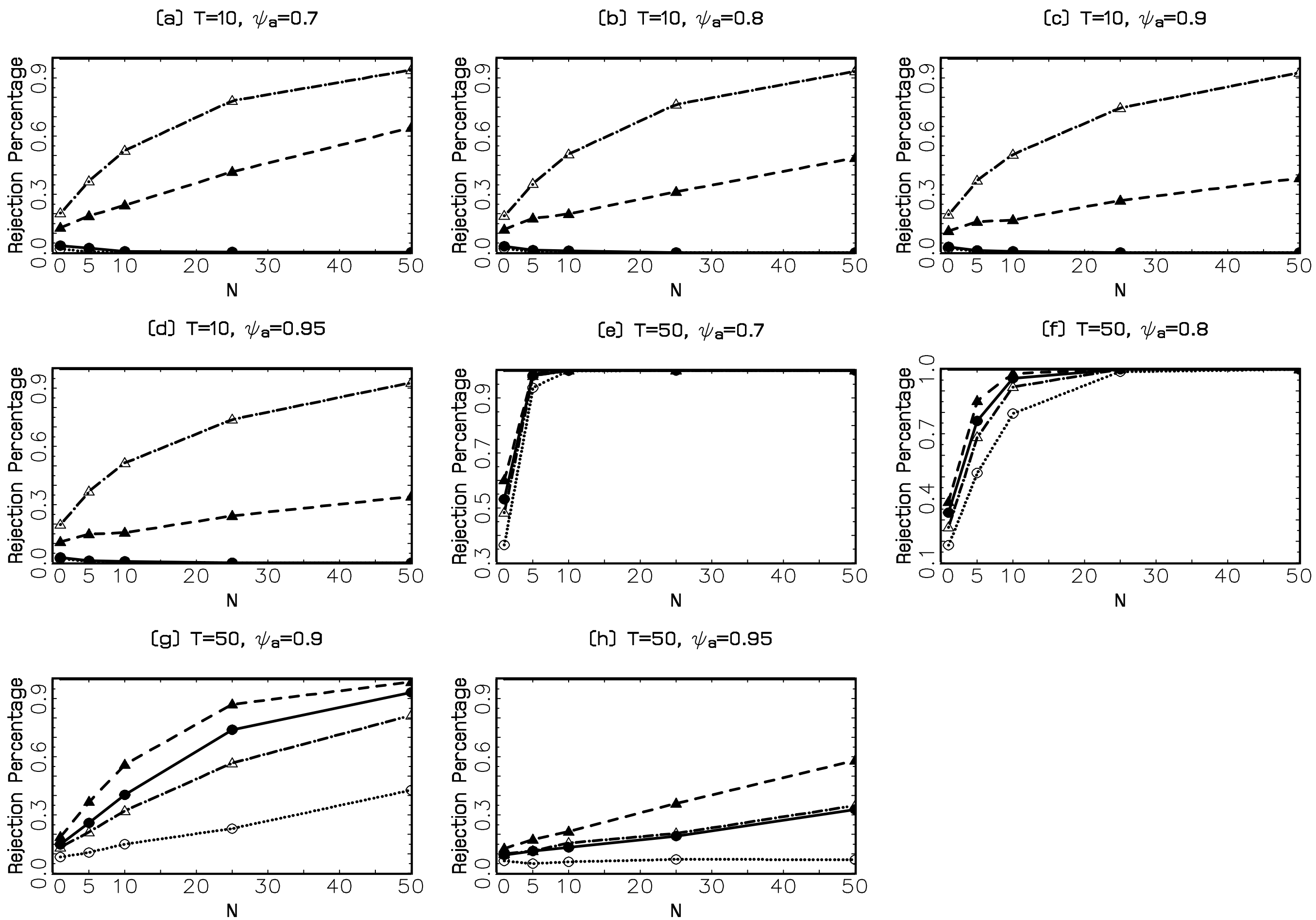


**Figure 5.3:** Small sample power results of the tests for DGP A and true cointegrating rank of two when hypothesized rank is one. $\bullet$——— panel SL-asym, $\blacktriangle$ − − − − panel SL-VAR(1), $\circ$........ LLL-asym, $\triangle$ − · − · −· LLL-VAR(1).

**Table 5.6:** Empirical size results of the tests for DGP B and true cointegrating rank of zero.

| | | panel SL | | LLL | |
|---|---|---|---|---|---|
| T-2 | N | asym | VAR(1) | asym | VAR(1) |
| 10 | 1 | 0.161 | 0.497 | 0.446 | 0.925 |
| | 5 | 0.277 | 0.898 | 0.905 | 1.000 |
| | 10 | 0.430 | 0.985 | 0.997 | 1.000 |
| | 25 | 0.694 | 1.000 | 1.000 | 1.000 |
| | 50 | 0.902 | 1.000 | 1.000 | 1.000 |
| 25 | 1 | 0.095 | 0.203 | 0.129 | 0.356 |
| | 5 | 0.105 | 0.386 | 0.238 | 0.746 |
| | 10 | 0.092 | 0.480 | 0.306 | 0.928 |
| | 25 | 0.147 | 0.782 | 0.596 | 0.998 |
| | 50 | 0.216 | 0.964 | 0.858 | 1.000 |
| 50 | 1 | 0.082 | 0.129 | 0.110 | 0.196 |
| | 5 | 0.074 | 0.170 | 0.121 | 0.325 |
| | 10 | 0.094 | 0.251 | 0.148 | 0.462 |
| | 25 | 0.095 | 0.349 | 0.240 | 0.756 |
| | 50 | 0.128 | 0.564 | 0.401 | 0.933 |
| 100 | 1 | 0.076 | 0.093 | 0.075 | 0.098 |
| | 5 | 0.073 | 0.110 | 0.086 | 0.158 |
| | 10 | 0.066 | 0.123 | 0.090 | 0.187 |
| | 25 | 0.067 | 0.164 | 0.141 | 0.316 |
| | 50 | 0.093 | 0.241 | 0.226 | 0.511 |
| 200 | 1 | 0.063 | 0.071 | 0.077 | 0.084 |
| | 5 | 0.057 | 0.070 | 0.068 | 0.087 |
| | 10 | 0.066 | 0.081 | 0.085 | 0.102 |
| | 25 | 0.079 | 0.110 | 0.132 | 0.162 |
| | 50 | 0.074 | 0.123 | 0.155 | 0.208 |
| 500 | 1 | 0.055 | 0.055 | 0.063 | 0.062 |
| | 5 | 0.062 | 0.063 | 0.071 | 0.070 |
| | 10 | 0.057 | 0.059 | 0.092 | 0.083 |
| | 25 | 0.063 | 0.064 | 0.090 | 0.082 |
| | 50 | 0.074 | 0.075 | 0.135 | 0.107 |
| 1000 | 1 | 0.063 | 0.063 | 0.061 | 0.056 |
| | 5 | 0.064 | 0.064 | 0.083 | 0.070 |
| | 10 | 0.065 | 0.065 | 0.073 | 0.068 |
| | 25 | 0.062 | 0.062 | 0.086 | 0.071 |
| | 50 | 0.078 | 0.078 | 0.111 | 0.077 |

**Table 5.7:** Empirical size results of the tests for DGP B and true cointegrating rank of one.

| | | $\psi = 0.5$ | | | | $\psi = 0.75$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | panel SL | | LLL | | panel SL | | LLL | |
| T-2 | N | asym | VAR(1) | asym | VAR(1) | asym | VAR(1) | asym | VAR(1) |
| 10 | 1 | 0.037 | 0.096 | 0.074 | 0.326 | 0.030 | 0.094 | 0.067 | 0.352 |
| | 5 | 0.013 | 0.106 | 0.074 | 0.650 | 0.014 | 0.109 | 0.069 | 0.698 |
| | 10 | 0.004 | 0.097 | 0.059 | 0.871 | 0.007 | 0.093 | 0.098 | 0.891 |
| | 25 | 0.001 | 0.090 | 0.084 | 0.992 | 0.000 | 0.078 | 0.117 | 0.998 |
| | 50 | 0.000 | 0.074 | 0.095 | 1.000 | 0.000 | 0.065 | 0.189 | 1.000 |
| 25 | 1 | 0.033 | 0.064 | 0.009 | 0.045 | 0.027 | 0.040 | 0.009 | 0.047 |
| | 5 | 0.012 | 0.045 | 0.000 | 0.044 | 0.006 | 0.020 | 0.005 | 0.034 |
| | 10 | 0.007 | 0.035 | 0.003 | 0.027 | 0.000 | 0.011 | 0.001 | 0.016 |
| | 25 | 0.002 | 0.037 | 0.001 | 0.028 | 0.000 | 0.002 | 0.000 | 0.015 |
| | 50 | 0.001 | 0.044 | 0.000 | 0.019 | 0.000 | 0.001 | 0.000 | 0.005 |
| 50 | 1 | 0.047 | 0.074 | 0.017 | 0.036 | 0.018 | 0.026 | 0.012 | 0.018 |
| | 5 | 0.034 | 0.059 | 0.008 | 0.031 | 0.004 | 0.007 | 0.000 | 0.004 |
| | 10 | 0.030 | 0.064 | 0.004 | 0.020 | 0.000 | 0.002 | 0.000 | 0.000 |
| | 25 | 0.014 | 0.072 | 0.000 | 0.012 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 50 | 0.028 | 0.104 | 0.001 | 0.010 | 0.000 | 0.000 | 0.000 | 0.000 |
| 100 | 1 | 0.078 | 0.092 | 0.050 | 0.072 | 0.016 | 0.019 | 0.005 | 0.012 |
| | 5 | 0.050 | 0.073 | 0.031 | 0.055 | 0.005 | 0.006 | 0.002 | 0.005 |
| | 10 | 0.058 | 0.094 | 0.023 | 0.043 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 25 | 0.059 | 0.110 | 0.027 | 0.058 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 50 | 0.087 | 0.137 | 0.027 | 0.067 | 0.000 | 0.000 | 0.000 | 0.000 |
| 200 | 1 | 0.074 | 0.077 | 0.073 | 0.084 | 0.031 | 0.033 | 0.012 | 0.014 |
| | 5 | 0.045 | 0.050 | 0.056 | 0.069 | 0.004 | 0.004 | 0.001 | 0.002 |
| | 10 | 0.075 | 0.082 | 0.075 | 0.090 | 0.003 | 0.004 | 0.000 | 0.001 |
| | 25 | 0.069 | 0.081 | 0.076 | 0.093 | 0.000 | 0.001 | 0.000 | 0.000 |
| | 50 | 0.074 | 0.090 | 0.074 | 0.106 | 0.000 | 0.000 | 0.000 | 0.000 |
| 500 | 1 | 0.051 | 0.064 | 0.060 | 0.061 | 0.048 | 0.043 | 0.031 | 0.033 |
| | 5 | 0.075 | 0.069 | 0.078 | 0.081 | 0.035 | 0.037 | 0.034 | 0.034 |
| | 10 | 0.074 | 0.088 | 0.069 | 0.069 | 0.030 | 0.025 | 0.025 | 0.025 |
| | 25 | 0.064 | 0.063 | 0.077 | 0.077 | 0.009 | 0.013 | 0.010 | 0.010 |
| | 50 | 0.068 | 0.084 | 0.075 | 0.073 | 0.004 | 0.008 | 0.013 | 0.013 |
| 1000 | 1 | 0.057 | 0.058 | 0.058 | 0.058 | 0.058 | 0.059 | 0.054 | 0.054 |
| | 5 | 0.065 | 0.065 | 0.076 | 0.075 | 0.046 | 0.046 | 0.073 | 0.073 |
| | 10 | 0.075 | 0.076 | 0.068 | 0.067 | 0.052 | 0.053 | 0.070 | 0.068 |
| | 25 | 0.072 | 0.075 | 0.064 | 0.061 | 0.038 | 0.040 | 0.055 | 0.052 |
| | 50 | 0.057 | 0.063 | 0.089 | 0.085 | 0.027 | 0.029 | 0.087 | 0.078 |

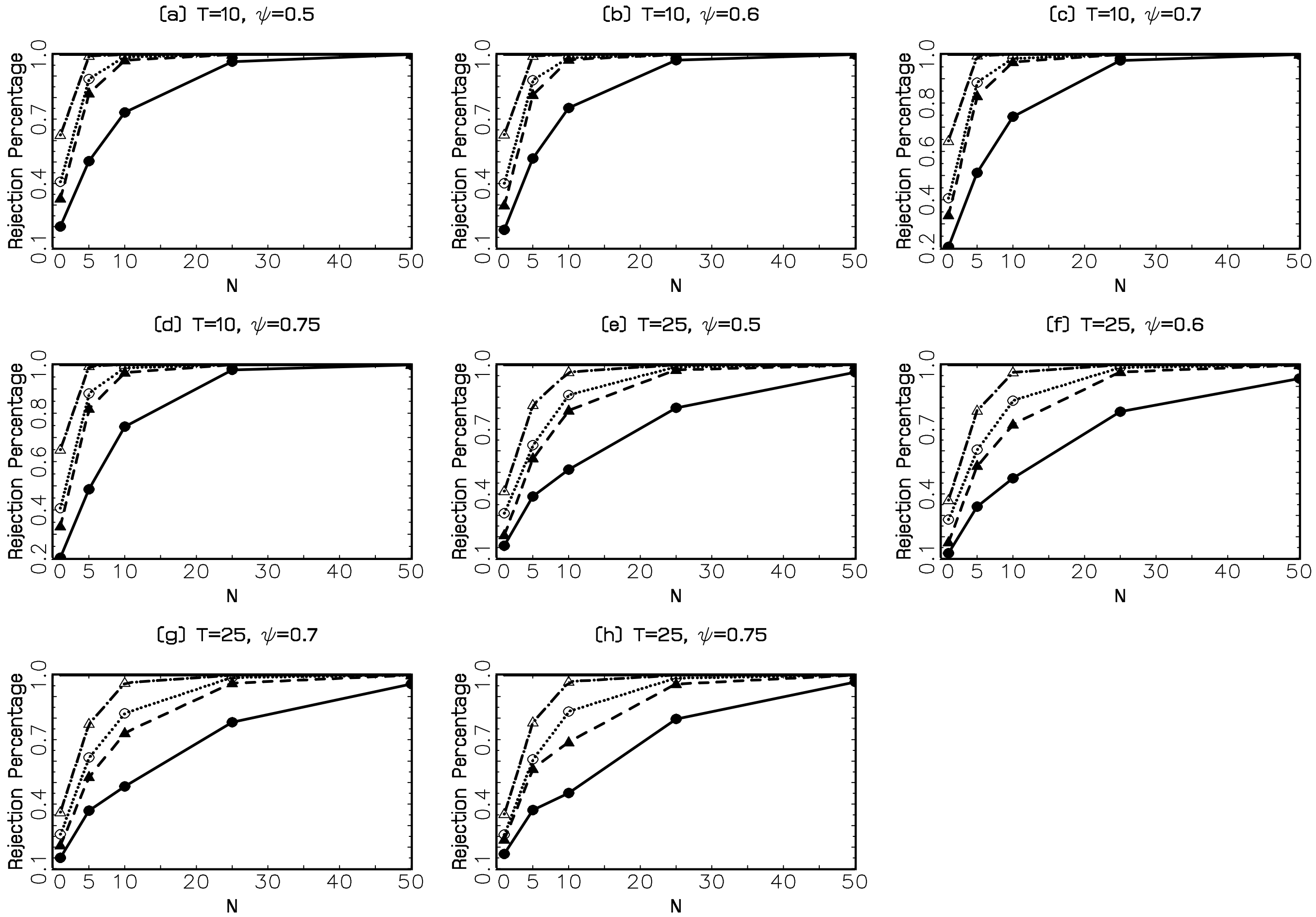


**Figure 5.4:** Small sample size-adjusted power results of the tests for DGP B and true cointegrating rank of one. ●——— panel SL-asym, ▲ − − − − panel SL-VAR(1), ○........ LLL-asym, △ − · − · −· LLL-VAR(1).

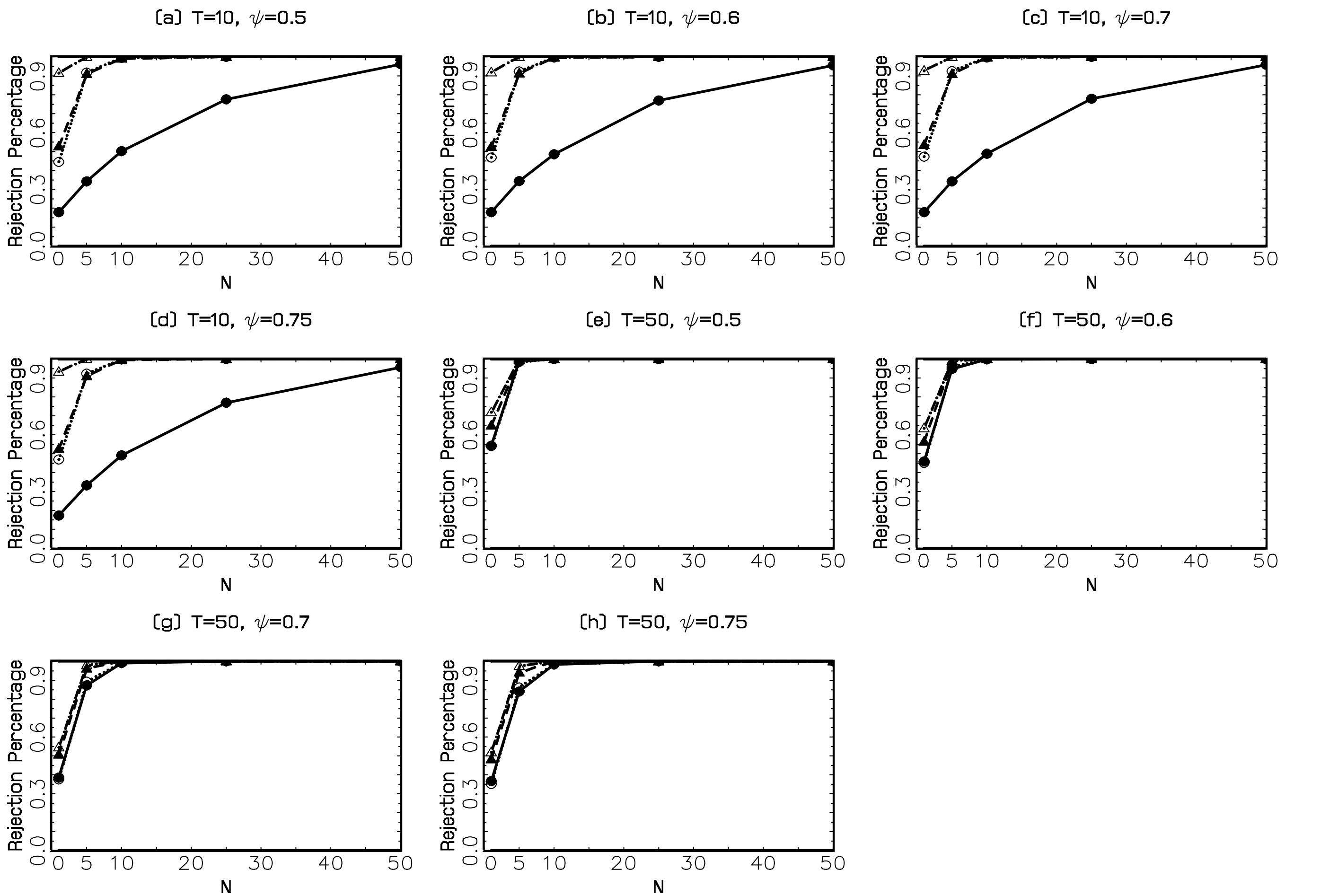


**Figure 5.5:** Small sample power results of the tests for DGP B and true cointegrating rank of two when hypothesized rank is zero. ●——— panel SL-asym, ▲ – – – – panel SL-VAR(1), ○........ LLL-asym, △ – · – · –· LLL-VAR(1).

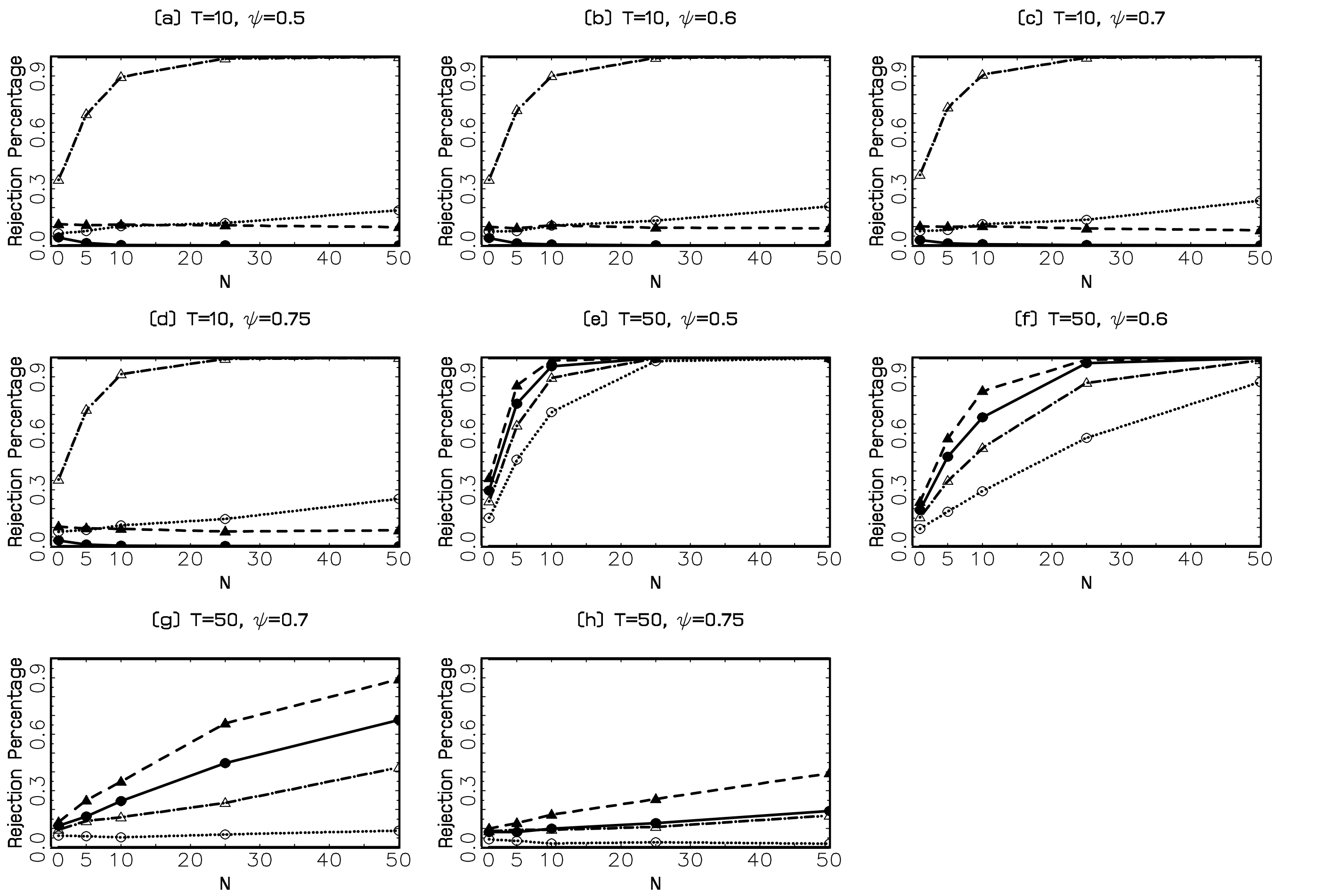


**Figure 5.6:** Small sample power results of the tests for DGP B and true cointegrating rank of two when hypothesized rank is one. ●——— panel SL-asym, ▲ – – – – panel SL-VAR(1), ○........ LLL-asym, △ – · – · –· LLL-VAR(1).

**Table 5.8:** Empirical size results of the tests for DGP C and true cointegrating rank of one.

| | | $\vartheta_i = 0$ | | | | $\vartheta_i \sim U(0, 0.5)$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | panel SL | | LLL | | panel SL | | LLL | |
| T-1 | N | asym | VAR(1) | asym | VAR(1) | asym | VAR(1) | asym | VAR(1) |
| 10 | 1 | 0.046 | 0.116 | 0.027 | 0.225 | 0.046 | 0.123 | 0.033 | 0.204 |
| | 5 | 0.010 | 0.090 | 0.007 | 0.412 | 0.011 | 0.083 | 0.005 | 0.320 |
| | 10 | 0.007 | 0.083 | 0.004 | 0.566 | 0.005 | 0.091 | 0.005 | 0.472 |
| | 25 | 0.000 | 0.065 | 0.002 | 0.816 | 0.000 | 0.062 | 0.000 | 0.671 |
| | 50 | 0.000 | 0.055 | 0.000 | 0.968 | 0.000 | 0.040 | 0.000 | 0.885 |
| 25 | 1 | 0.047 | 0.080 | 0.045 | 0.099 | 0.044 | 0.074 | 0.027 | 0.073 |
| | 5 | 0.018 | 0.064 | 0.024 | 0.142 | 0.020 | 0.066 | 0.006 | 0.035 |
| | 10 | 0.016 | 0.072 | 0.015 | 0.155 | 0.021 | 0.081 | 0.002 | 0.042 |
| | 25 | 0.012 | 0.082 | 0.011 | 0.236 | 0.011 | 0.072 | 0.000 | 0.025 |
| | 50 | 0.004 | 0.095 | 0.007 | 0.347 | 0.002 | 0.062 | 0.000 | 0.015 |
| 50 | 1 | 0.043 | 0.061 | 0.067 | 0.106 | 0.077 | 0.105 | 0.037 | 0.061 |
| | 5 | 0.053 | 0.089 | 0.053 | 0.121 | 0.069 | 0.104 | 0.013 | 0.039 |
| | 10 | 0.045 | 0.086 | 0.056 | 0.139 | 0.059 | 0.115 | 0.011 | 0.028 |
| | 25 | 0.020 | 0.070 | 0.036 | 0.156 | 0.059 | 0.126 | 0.001 | 0.019 |
| | 50 | 0.024 | 0.093 | 0.041 | 0.218 | 0.054 | 0.173 | 0.000 | 0.002 |
| 100 | 1 | 0.060 | 0.074 | 0.060 | 0.074 | 0.076 | 0.089 | 0.047 | 0.056 |
| | 5 | 0.065 | 0.078 | 0.068 | 0.098 | 0.078 | 0.112 | 0.021 | 0.032 |
| | 10 | 0.053 | 0.087 | 0.066 | 0.115 | 0.108 | 0.140 | 0.009 | 0.019 |
| | 25 | 0.042 | 0.073 | 0.055 | 0.109 | 0.108 | 0.174 | 0.003 | 0.008 |
| | 50 | 0.039 | 0.079 | 0.063 | 0.149 | 0.149 | 0.250 | 0.001 | 0.006 |
| 200 | 1 | 0.073 | 0.077 | 0.091 | 0.099 | 0.082 | 0.085 | 0.050 | 0.058 |
| | 5 | 0.074 | 0.082 | 0.071 | 0.089 | 0.107 | 0.117 | 0.024 | 0.027 |
| | 10 | 0.055 | 0.063 | 0.078 | 0.093 | 0.108 | 0.120 | 0.012 | 0.016 |
| | 25 | 0.065 | 0.074 | 0.071 | 0.099 | 0.150 | 0.170 | 0.006 | 0.011 |
| | 50 | 0.052 | 0.064 | 0.074 | 0.106 | 0.197 | 0.232 | 0.001 | 0.001 |
| 500 | 1 | 0.069 | 0.069 | 0.055 | 0.056 | 0.090 | 0.091 | 0.057 | 0.057 |
| | 5 | 0.073 | 0.073 | 0.064 | 0.065 | 0.115 | 0.116 | 0.028 | 0.029 |
| | 10 | 0.058 | 0.059 | 0.058 | 0.059 | 0.107 | 0.108 | 0.019 | 0.019 |
| | 25 | 0.060 | 0.063 | 0.057 | 0.057 | 0.151 | 0.160 | 0.007 | 0.007 |
| | 50 | 0.061 | 0.062 | 0.081 | 0.079 | 0.242 | 0.255 | 0.000 | 0.000 |
| 1000 | 1 | 0.074 | 0.075 | 0.073 | 0.073 | 0.089 | 0.089 | 0.044 | 0.044 |
| | 5 | 0.064 | 0.064 | 0.055 | 0.054 | 0.129 | 0.129 | 0.018 | 0.017 |
| | 10 | 0.067 | 0.067 | 0.061 | 0.059 | 0.135 | 0.139 | 0.011 | 0.011 |
| | 25 | 0.069 | 0.072 | 0.073 | 0.072 | 0.183 | 0.191 | 0.002 | 0.002 |
| | 50 | 0.069 | 0.075 | 0.080 | 0.074 | 0.259 | 0.269 | 0.001 | 0.001 |

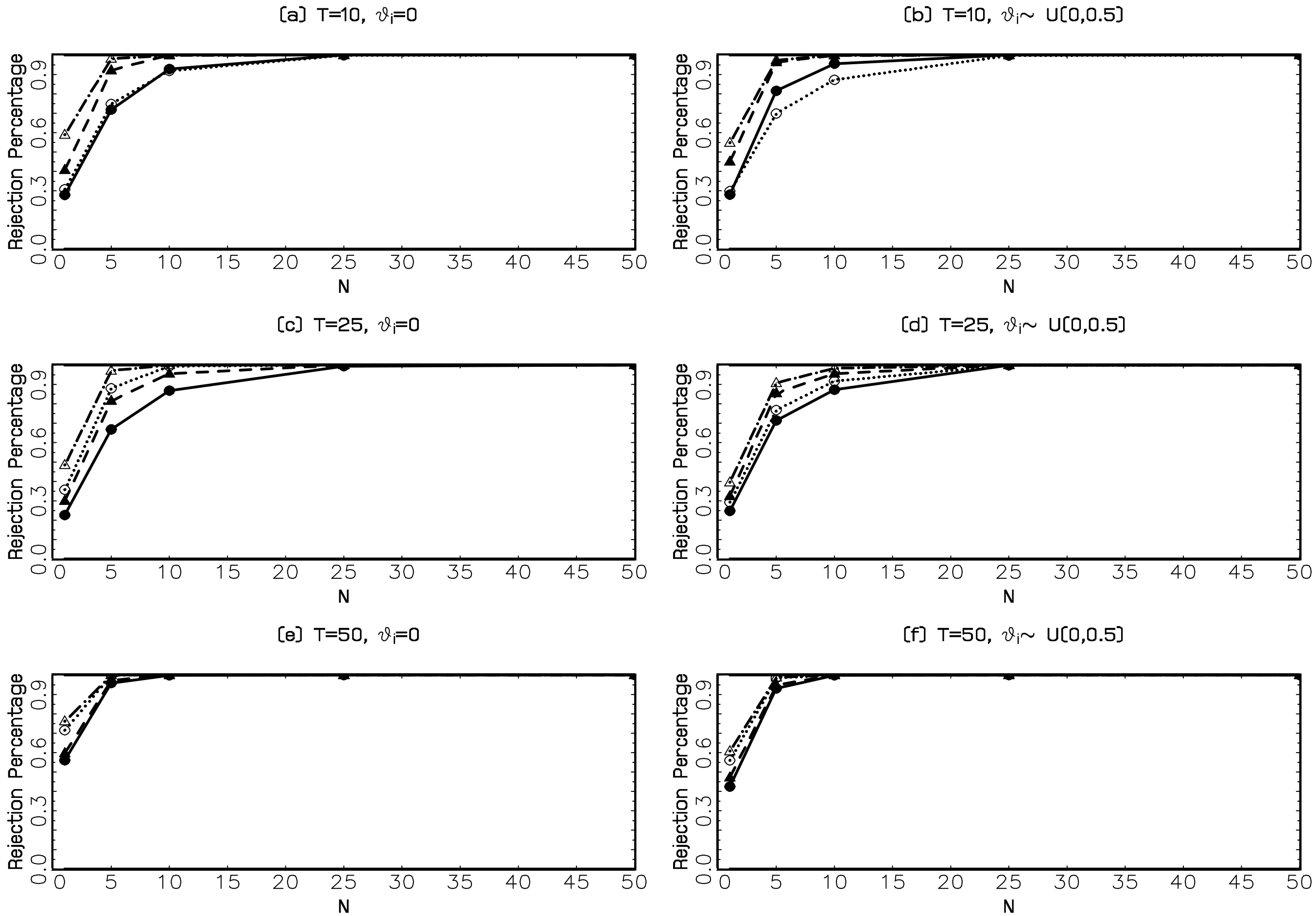


**Figure 5.7:** Small sample size-adjusted power results of the tests for DGP C and true cointegrating rank of one. •——— panel SL-asym, ▲ − − − − panel SL-VAR(1), ○........ LLL-asym, △ − · − · −· LLL-VAR(1).

In general, with VAR(1) moments the panel SL test has better size properties for $T \leq 50$, and with asymptotic moments the test exhibits a reasonable size for $T \geq 100$. With the increase in $T$ once more the size results of the tests based on two different approximations converge to each other. It is apparent from Table 5.7 that both tests are undersized when $\psi$ increases from 0.5 to 0.75, whereas for $T = 1000$ the sizes of the tests converge to the 5% nominal level. When the test statistics are approximated with VAR(1) moments the tests are undersized if $T \leq 500$, except for $T = 10$, and their sizes approach the 5% level for $T = 1000$.

Figures 5.4-5.6 display the size-adjusted power results for DGP B. The size-adjusted powers of both tests for the true cointegrating rank of one approach unity with the increment in $N$ even for small $T$. Moreover, the power of the LLL test is higher than the power of the panel SL test, which is most obvious for $T = 10$. Once more the approximation with VAR(1) moments delivers higher power than the approximation with asymptotic moments.

In Figure 5.6 it is depicted that the power of the LLL test is higher than the power of the panel SL test for $T = 10$. The false hypothesis of one cointegrating relation cannot be rejected for the panel SL test when it is based on asymptotic moments. In addition to this, if VAR(1) moments are used the power of the LLL test for $T = 10$ approaches unity, which is not the case for the panel SL test. On the contrary, the panel SL test shows better power than the LLL test with an increase in $T$ to 50. In addition to this, the power of both tests decreases if $\psi$ increases, which is in line with the simulation results of DGP A.

Hence, for DGP B we can conclude that the panel SL test shows better size properties than the LLL test. As outlined above the power of the LLL test based on the approximation with VAR(1) moments is the highest among the considered tests and approximations.

### Simulation Results for DGP C

If there is no correlation and the test statistics are approximated with asymptotic moments, both panel tests are undersized for small $T$ and their sizes are around the 5% level for large $T$ (see Table 5.8). Based on VAR(1) moments, the size of the panel SL test ranges from 0.055 (for $T = 10$, $N = 50$) to 0.095 (for $T = 25$, $N = 50$), especially for $N \leq 5$. On the contrary, if the LLL test statistic is approximated with VAR(1) moments, the test is oversized for $T \leq 200$, and its size is close to the 5% level for $T \geq 500$.

Based on the approximation with asymptotic moments, the panel SL test has slightly better size properties than the LLL test for $T = 100, 200$.

In addition to this, if the asymptotic moments are used and there is correlation between the components of the DGP, the panel SL test is undersized for $T = 10$, 25 and it becomes oversized with an increase in $T$ and $N$, e.g. 0.259 (for $T = 1000$ and $N = 50$). However, the size of the panel SL test is 0.054 for $T = 50$, $N = 50$. If asymptotic moments are used, the LLL test is undersized for almost all $T$ values. Furthermore, based on VAR(1) moments for almost all combinations of $T$ and $N$, the panel SL test is oversized, whereas the LLL test is just oversized if $T = 10$, and it becomes undersized as $T$ and $N$ rise. Thus, the size of the LLL test does not approach the 5% level, except for $T \geq 50$ and $N = 1$. However, then the LLL test is just the standardized version of the multivariate Johansen trace test which allows a linear time trend in the data.

The size-adjusted power results are similar for both tests, independent of which approximation method is used. The power of the tests converge to unity with an increase in $N$, even for small $T$. This means that the probability of rejecting the false hypothesis of no cointegrating relation is one. Here we just illustrate the size-adjusted power results for $T = 10, 25, 50$ because if $T \geq 50$, the powers of the tests converge to unity even for $N = 1$. The panel SL test has slightly less power than the LLL test, but the difference disappears as $T$ rises.

For DGP C the panel SL test has again the best size properties. Both tests are size distorted when there is correlation between the components of the process. Hence, the power of the LLL test is slightly higher than the panel SL test.

## 5.5 Conclusions

In this chapter a new maximum-likelihood-based panel cointegration test (i.e. the panel SL test) was introduced. It allows for a linear time trend in the DGP and is an extension of the multivariate cointegration test ($\mathrm{LR}_{\mathrm{trace}}^{\mathrm{GLS}}$ test) of Saikkonen and Lütkepohl (2000a). To find out the finite sample properties of the panel SL test, in a Monte Carlo study three different DGPs were considered and the results were compared with the Larsson et al. (2001) test (i.e. the LLL test), which allows a linear time trend in the data.

The simulation results indicate size distortions for small $T$. The sizes of both tests come close to the nominal 5% rejection level as $T$ increases. In general the panel SL test has better size properties than the LLL test, especially if there is no correlation between the components of the DGP. Also for small $T$, if VAR(1) moments are used the panel SL test delivers better size properties in comparison to the LLL test, which is severely oversized for small $T$ independent of the approximation chosen. Moreover, the sizes of both tests with different approximations converge to each other with an increase in $T$.

With the introduction of correlation between stationary and nonstationary components of the processes, size distortions are depicted, however the panel SL test has still reasonable size for large $T$. In addition to this, we found out that, if the DGP consists of a near nonstationary component, then the tests become size distorted and lose power.

In general, the powers of both panel cointegration tests approach unity if $N$ increases for small $T$. Additionally, the approximation based on VAR(1) moments delivers tests with higher power than the approximation based on asymptotic moments.

## 5.6 Mathematical Appendix

Here we demonstrate the proof of Lemma 5.1 for $d = 1$. We can rewrite Lemma 5.1 for $d = 1$ as:

**Lemma 5.2:** *There are some constants $a$ and $b$ such that for all $T$,*

(i.) $E(\widetilde{Z}^2_{T,1}) < a$,

(ii.) $E(\widetilde{Z}^4_{T,1}) < b$.

*Proof.* This proof is analogous to the proof of Lemma 4.1. Note that the $\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}}$ statistic is

$$\widetilde{Z}_{T,d} = \mathrm{tr}(B_T' A_T^{-1} B_T) \xrightarrow{w} \widetilde{Z}_d. \tag{5.28}$$

$A_T$ and $B_T$ are defined as in (5.18) and (5.19), respectively. $\widetilde{Z}_d$ denotes the asymptotic distribution of the $\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}}$ statistic. For simplicity, we assume that $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)'$ and $\varepsilon_t \sim N(0,1)$ *i.i.d.* If we let $\widetilde{\varepsilon}_t = \varepsilon_t - \overline{\varepsilon}$ for $t = 1, \ldots, T$ with $\overline{\varepsilon} = T^{-1} \sum_{t=1}^T \varepsilon_t$, then the trace statistic $\widetilde{Z}_{T,d}$ can be rewritten as

$$\widetilde{Z}_{T,d} \quad = \quad \mathrm{tr}(B_T' A_T^{-1} B_T) = \mathrm{tr}(\widetilde{\varepsilon}' \widetilde{P}_{\widetilde{Y}} \widetilde{\varepsilon}), \tag{5.29}$$

in which $\widetilde{Y} = BA\widetilde{\varepsilon}$ and

$$\widetilde{P}_{\widetilde{Y}} \quad = \quad BA\widetilde{\varepsilon}(\widetilde{\varepsilon}' A' B' BA\widetilde{\varepsilon})^{-1} \widetilde{\varepsilon}' A' B', \tag{5.30}$$

is the projection matrix onto the column space of $\widetilde{Y}$. Note that $A$ and $B$ matrices are defined as in Chapter 4 and

$$\widetilde{\varepsilon} \quad = \quad \begin{pmatrix} \varepsilon_1 - \bar{\varepsilon} \\ \vdots \\ \varepsilon_T - \bar{\varepsilon} \end{pmatrix} = \varepsilon - (1\!\!1_T \otimes \bar{\varepsilon}) = (I_T - P)\varepsilon = Q\varepsilon, \tag{5.31}$$

with $\bar{\varepsilon} = \frac{1}{T}\mathbb{1}_T'\varepsilon$ and $\mathbb{1}_T$ denoting a $T$-dimensional column vector of ones. The symmetric and idempotent matrices $P$ and $Q$ are defined as in Section 4.3. Furthermore,

$$\widetilde{Z}_{T,d} \;=\; \operatorname{tr}(\widetilde{\varepsilon}'\widetilde{P}_{\widetilde{Y}}\widetilde{\varepsilon}) \le \operatorname{tr}(\widetilde{\varepsilon}'\widetilde{\varepsilon}) = \operatorname{tr}(\varepsilon' Q\varepsilon) \le \operatorname{tr}(\varepsilon'\varepsilon) \tag{5.32}$$

because $(I_T - \widetilde{P}_{\widetilde{Y}})$ and $(I_T - Q)$ are nonnegative definite matrices. Consequently, $\widetilde{\varepsilon}'(I_T - \widetilde{P}_{\widetilde{Y}})\widetilde{\varepsilon} = \widetilde{\varepsilon}'\widetilde{\varepsilon} - \widetilde{\varepsilon}'\widetilde{P}_{\widetilde{Y}}\widetilde{\varepsilon}$ and $\varepsilon'(I_T - Q)\varepsilon = \varepsilon'\varepsilon - \varepsilon' Q\varepsilon$ are nonnegative definite, too. Since $\varepsilon'\varepsilon \sim W_d(T, \Omega)$, all moments of $\widetilde{Z}_{T,d}$ exist because of the same reasons discussed in Section 4.2. In other words, the moments of $Z_{T,1}$ exist for all $T$. Thus,

$$\begin{aligned} \widetilde{Z}_{T,d} &= \varepsilon' QBAQ\varepsilon\left(\varepsilon' QA'B'BAQ\varepsilon\right)^{-1}\varepsilon' QA'B'Q\varepsilon \\ &= \varepsilon' QDQ\varepsilon\left(\varepsilon' QD'DQ\varepsilon\right)^{-1}\varepsilon' QD'Q\varepsilon, \end{aligned} \tag{5.33}$$

with $D = BA$. To prove (i.) of Lemma 5.2, first we rewrite $\widetilde{Z}_{T,d}$ for $d = 1$.

$$\widetilde{Z}_{T,1} = \frac{\left(\varepsilon' QDQ\varepsilon\right)^2}{\varepsilon' QD'DQ\varepsilon} \tag{5.34}$$

We already know that $\varepsilon' QDQ\varepsilon = \varepsilon' QSQ\varepsilon$, in which the matrix $S$ is defined as in (4.26). This leads to

$$R = QSQ = -\frac{1}{2}Q. \tag{5.35}$$

From (5.35) it is obvious that $R$ is a negative semidefinite matrix and has $(T-1)$ eigenvalues equal to $-\frac{1}{2}$ and its largest eigenvalue is zero. We can represent (5.34) using (5.35) also as

$$\widetilde{Z}_{T,1} = \frac{\left(\varepsilon' R\varepsilon\right)^2}{\varepsilon' H\varepsilon}, \tag{5.36}$$

with $H = QD'DQ = QFQ$ and $F = D'D$. To prove $E(\widetilde{Z}_{T,1}) < a^*$, for all $T$ and $a^* \in (0, \infty)$, we use the Cauchy-Schwarz inequality.

$$E(\widetilde{Z}_{T,1}) = E\left[\frac{\left(\varepsilon' R\varepsilon\right)^2}{\varepsilon' H\varepsilon}\right] \le \sqrt{E\left[\left(\varepsilon' R\varepsilon\right)^4\right] E\left[\frac{1}{\left(\varepsilon' H\varepsilon\right)^2}\right]} \tag{5.37}$$

Note that $\varepsilon' R\varepsilon = -\frac{1}{2}\zeta_1$, and $\zeta_1 = \chi^2_{(T-1)}$. Using the fourth moment of the $\chi^2$ distribution with $n$ degrees of freedom, it can be shown that

$$\begin{aligned} E\left[\left(\varepsilon' R\varepsilon\right)^4\right] &= \frac{1}{16}E(\zeta_1^4) && (5.38) \\ &= \frac{1}{16}(T+5)(T+3)(T+1)(T-1) && (5.39) \\ &= c_1 T^4 + o(T^4), && (5.40) \end{aligned}$$

for $c_1 \in (0, \infty)$.

The spectral decomposition representation of the matrix $H$ is denoted by

$$H = V'\Lambda V = \sum_{t=1}^{T} \lambda_t \mathrm{v}_t \mathrm{v}_t'. \tag{5.41}$$

This time $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_T)$ is the diagonal matrix consisting of the eigenvalues of the matrix $H$, and $V = (\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_T)$ is the $(T \times T)$ orthogonal matrix of the eigenvectors corresponding to the eigenvalues of the matrix $H$. Since $\xi_t = \varepsilon' \mathrm{v}_t \sim N(0,1)$ *i.i.d.* with $\mathrm{v}_t'\mathrm{v}_s = \begin{cases} 1, & \text{for} \quad t = s \\ 0, & \text{for} \quad t \neq s \end{cases}$, and $\varepsilon \sim N(0, I_T)$,

$$\varepsilon' H \varepsilon = \sum_{t=1}^{T} \lambda_t \varepsilon' \mathrm{v}_t \mathrm{v}_t' \varepsilon = \sum_{t=1}^{T} \lambda_t \xi_t^2, \tag{5.42}$$

in which $\xi_t^2 \sim \chi^2_{(1)}$ for $t = 1, \ldots, T$, and the $\xi_t^2$'s are mutually independently distributed.

The eigenvalues of the positive semidefinite matrix $H = QFQ$ are[8]

$$\lambda_t = \frac{1}{2 - 2\cos\left(\frac{t\pi}{T}\right)} \quad \text{for} \quad t = 1, \ldots, T-1, \quad \lambda_T = 0, \tag{5.43}$$

with $\lambda_1 > \lambda_2 > \ldots > \lambda_{T-1} > \lambda_T$. Moreover, $\lambda_1 = \lambda_{\max} = c_2 T^2 + o(T^2)$ for some $c_2 \in (0, \infty)$ and $\lim\limits_{T \to \infty} \lambda_{T-1} = \frac{1}{4}$. Analogous to the proof in Chapter 4, the following inequality will help us to achieve an upper bound for $E[(\varepsilon' H \varepsilon)^{-2}]$.

$$E\left[\frac{1}{(\varepsilon' H \varepsilon)^2}\right] \leq \frac{1}{\lambda_5^2} E\left[\frac{1}{(z_1)^2}\right], \tag{5.44}$$

in which $z_1 \sim \chi^2_{(5)}$ and $\lambda_5$ is the fifth largest eigenvalue of $H$. On account of (5.37), (5.40) and (5.44) we obtain

$$E(\widetilde{Z}_{T,1}) \leq \sqrt{E\left[(\varepsilon' R \varepsilon)^4\right] \frac{1}{\lambda_5^2} E\left[\frac{1}{(z_1)^2}\right]} = \sqrt{\frac{1}{3}[c_1 T^4 + o(T^4)] \frac{1}{\lambda_5^2}}, \tag{5.45}$$

[8]The positive eigenvalues of the matrix $H$ are the same as the eigenvalues of the inverse of the following tridiagonal matrix:

$$G = (-1) \begin{pmatrix} -2 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & -2 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & -2 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1 & -2 \end{pmatrix}.$$

The eigenvalues of such a matrix are discussed in Yueh (2005).

with $E[(z_1)^{-2}] = \frac{1}{3}$. As a result, to prove the uniform boundedness of $E(\widetilde{Z}_{T,1})$, it is necessary to show that $\lambda_5 = c_3T^2 + o(T^2)$ for $c_3 \in (0, \infty)$. In other words;

$$\frac{\lambda_5}{T^2} = \frac{1}{2\left[1 - \cos\left(\frac{5\pi}{T}\right)\right]T^2} \to c_3 \quad \text{as } T \to \infty \text{ with } c_3 \in (0, \infty). \tag{5.46}$$

Let

$$f(T) = T^2\left[1 - \cos\left(\frac{5\pi}{T}\right)\right]. \tag{5.47}$$

If it can be shown that $f(T) \to c_1^*$ as $T \to \infty$ with $c_1^* \in (0, \infty)$, this will complete the proof of $E(\widetilde{Z}_{T,1}) < a^*$ for all $T$. Hence,

$$\lim_{T\to\infty} f(T) = \lim_{T\to\infty} \frac{1 - \cos\left(\frac{5\pi}{T}\right)}{\frac{1}{T^2}} = \lim_{T\to\infty} \frac{g(T)}{h(T)} \left( = \frac{0\text{''}}{\text{``}0}\right) \tag{5.48}$$

as $\lim\limits_{T\to\infty} \cos\left(\frac{5\pi}{T}\right) = 1$ and $\lim\limits_{T\to\infty} \frac{1}{T^2} = 0$. By applying l'Hospital rule twice we obtain

$$\lim_{T\to\infty} f(T) = \frac{(5\pi)^2}{2} > 0, \tag{5.49}$$

which leads to $\lambda_5 = c_3T^2 + o(T^2)$. Subsequently, $1/\lambda_5^2 = 1/[c_4T^4 + o(T^4)]$ for $c_4 \in (0, \infty)$. Finally,

$$\begin{aligned} E(\widetilde{Z}_{T,1}) &\leq \sqrt{E\left[(\varepsilon' R\varepsilon)^4\right] \frac{1}{\lambda_5^2} E\left[\frac{1}{(z_1)^2}\right]} && (5.50)\\ &= \sqrt{\frac{1}{3}\frac{[c_1T^4 + o(T^4)]}{[c_4T^4 + o(T^4)]}} \underset{T\to\infty}{\longrightarrow} \sqrt{\frac{c_1}{3c_4}}, && (5.51) \end{aligned}$$

and this completes the proof of $E(\widetilde{Z}_{T,1}) < a^*$ for all $T$ with some positive and finite $a^*$.

In order to prove $E(\widetilde{Z}_{T,1}^2) < a$ for all $T$, we apply again the Cauchy-Schwarz inequality, which leads to the following expression.

$$E(\widetilde{Z}_{T,1}^2) = E\left[\left(\frac{(\varepsilon' R\varepsilon)^2}{\varepsilon' H\varepsilon}\right)^2\right] \leq \sqrt{E[(\varepsilon' R\varepsilon)^8] E\left[\frac{1}{(\varepsilon' H\varepsilon)^4}\right]} \tag{5.52}$$

Hence, with the eighth moment of the $\chi^2$ distribution with $n$ degrees of freedom,

$$\begin{aligned} E\left[(\varepsilon' R\varepsilon)^8\right] &= \frac{1}{2^8} E[\zeta_1^8] && (5.53)\\ &= \frac{1}{256}(T+13)(T+11)(T+9)(T+7)(T+5) \\ &\quad (T+3)(T+1)(T-1) && (5.54)\\ &= c_5T^8 + o(T^8), \quad \text{with } c_5 \in (0, \infty). && (5.55) \end{aligned}$$

Applying a similar procedure as explained above, (i.) can be proven with the help of the following inequality.

$$E\left[\frac{1}{(\varepsilon' H\varepsilon)^4}\right] \leq \frac{1}{\lambda_9^4} E\left[\frac{1}{(z_2)^4}\right], \tag{5.56}$$

in which $z_2 \sim \chi^2_{(9)}$ and $\lambda_9$ is the ninth largest eigenvalue of $H$. Inserting (5.55) and (5.56) into (5.52), we obtain

$$\begin{aligned} E(\widetilde{Z}^2_{T,1}) &\leq \sqrt{E[(\varepsilon' R\varepsilon)^8]\frac{1}{\lambda_9^4}E\left[\frac{1}{(z_2)^4}\right]} & (5.57)\\ &= \sqrt{\frac{1}{105}[c_5T^8 + o(T^8)]\frac{1}{\lambda_9^4}}, & (5.58) \end{aligned}$$

with $E[(z_2)^{-4}] = 1/105$, which is finite and independent of $T$. In other words, to complete the proof of (i.), it should be demonstrated that

$$\frac{\lambda_9}{T^2} = \frac{1}{2\left[1-\cos\left(\frac{9\pi}{T}\right)\right]T^2} \to c_6 \quad \text{as } T\to\infty \text{ with } c_6 \in (0,\infty), \tag{5.59}$$

which can also be shown by

$$f(T) = T^2\left[1-\cos\left(\frac{9\pi}{T}\right)\right] \to c_2^* \quad \text{as } T\to\infty \text{ with } c_2^* \in (0,\infty). \tag{5.60}$$

Please note that

$$\lim_{T\to\infty} f(T) = \lim_{T\to\infty}\frac{1-\cos\left(\frac{9\pi}{T}\right)}{\frac{1}{T^2}} = \lim_{T\to\infty}\frac{g(T)}{h(T)}\left(= \frac{0"}{"0}\right). \tag{5.61}$$

By applying l'Hospital rule twice we get

$$\lim_{T\to\infty} f(T) = \frac{(9\pi)^2}{2} > 0. \tag{5.62}$$

Consequently, $\lambda_9 = c_6T^2 + o(T^2)$, or $1/\lambda_9^4 = 1/[c_7T^8 + o(T^8)]$, with $c_7 \in (0,\infty)$ This completes the proof of (i.) because

$$\begin{aligned} E(\widetilde{Z}^2_{T,1}) &\leq \sqrt{E\left[(\varepsilon' R\varepsilon)^8\right]\frac{1}{\lambda_9^4}E\left[\frac{1}{(z_2)^4}\right]} & (5.63)\\ &= \sqrt{\frac{1}{105}\frac{[c_5T^8 + o(T^8)]}{[c_7T^8 + o(T^8)]}} \underset{T\to\infty}{\longrightarrow} \sqrt{\frac{c_5}{105c_7}}. & (5.64) \end{aligned}$$

Thus, the second moments of $\widetilde{Z}_{T,1}$ are uniformly bounded as $E(\widetilde{Z}_{T,1}^2) < a$ for all $T$ with some positive and finite $a$.

The proof of (ii.) is analogous to the proof of (i.), and omitted here to save space.

**Theorem 5.3:** It holds that $E(\widetilde{Z}_{T,1}^2) < \infty$, $E(\widetilde{Z}_{T,1}^2) \to E(\widetilde{Z}_1^2)$ and $E(\widetilde{Z}_{T,1}) \to E(\widetilde{Z}_1)$ *as* $T \to \infty$.

*Proof.* Using the information that $\widetilde{Z}_{T,1}$ converges in distribution to $\widetilde{Z}_1$ (see Saikkonen and Lütkepohl, 2000a), Theorem 5.3 can be verified. The second moment of $\widetilde{Z}_1$ exists (with $E(\widetilde{Z}_{T,1}^2) \to E(\widetilde{Z}_1^2)$) if $\{\widetilde{Z}_{T,1}^2\}$ is uniformly integrable (see Serfling, 1980). A sufficient condition for the uniform integrability of $\{\widetilde{Z}_{T,1}^2\}$ is that $E|\widetilde{Z}_{T,1}|^{2+\delta}$ is uniformly bounded for some $\delta > 0$, i.e $\sup_T E|\widetilde{Z}_{T,1}|^{2+\delta} < \infty$. On account of Lemma 5.2, $E(\widetilde{Z}_1^2)$ exists as $E(\widetilde{Z}_{T,1}^2) < a < \infty$ and $E(\widetilde{Z}_{T,1}^4) < b < \infty$ for all $T$, which completes the proof.

# Chapter 6

# Money Demand: Evidence from OECD Countries

To put the panel SL test proposed in Chapter 5 into practice, in this chapter the money demand function is analyzed for a panel data of OECD countries. Money demand plays a key role in the determination of the monetary policy of central banks. Using unit root and cointegration techniques, the stability of the money demand is tested by examining the relation between money demand and some crucial macroeconomic variables such as income, interest rate etc. There is an extensive empirical literature on money demand functions based on the conventional cointegration techniques. The findings of most of these studies are summarized in Sriram (2001).

Unfortunately, as emphasized in Chapter 1 the conventional unit root and cointegration tests have low power against the stationary alternatives. Thus, long time series is necessary to increase the power of the conventional unit root and cointegration tests. Yet, there is not always enough observations for the variables under consideration. To overcome these problems the data can be extended using the information from different countries, which turns the time series framework into a panel data framework. In this way, the number of available observations is increased. Thus, to test the existence of a long-run stable money demand, the panel unit root and cointegration techniques can be used, which will provide gain in power.

There are also several empirical studies on money demand in panel cointegration literature. One of these important studies is from Mark and Sul (2003). In their study, they proposed the simple panel DOLS estimator which delivers more precise estimates than the single-equation estimator. With this estimator Mark and Sul (2003) estimated the M1 demand for a panel consisting of 19 OECD countries. Mark and Sul (2003) obtained an income elasticity near 1, and an interest rate semi-elasticity of -0.02. Considering a panel dataset consisting of six Gulf Cooperation Council (GCC) countries Harb (2004) tested the M1 demand using Pedroni (1999)'s panel cointegration tests and estimated the cointegrating equation with the FMOLS estimator developed by Pedroni (2000). Harb (2004) found a significant effect of the interest rate on the money demand. Furthermore, in another study Dreger et al. (2007) analyzed the broad money demand for 10 new European Union (EU) countries. Their income elasticity estimate is around 1.70, and interest rate semi-elasticity is negative. Fidrmuc (2008) analyzed M2 demand for a panel of six Central and Eastern European countries which are getting prepared to enter the European Economic and Monetary Union (EMU). Fidrmuc (2008) estimated the money demand equation both with panel FMOLS and DOLS estimators and concluded that the euro area interest rates have a significant effect on the money demand of these six countries. Moreover, with an unbalanced panel consisting of 15 Latin American countries, Carrera (2006) estimated the money demand function by the FMOLS estimator of Pedroni (2000) and found evidence for a cointegrating relation.

To analyze the long-run money demand relation, first the integratedness of the variables are detected by means of the panel unit root tests of Levin et al. (2002) and Im et al. (2003). Then, the existence of the number of cointegrating relations is checked by the panel SL and the residual-based tests of Pedroni (1999). Finally, the long-run money demand equation is estimated with the panel DOLS approach of Mark and Sul (2003).

This chapter is structured as follows: In Section 6.1 the economic theory behind the money demand function is explained. The data used for the empirical analysis is introduced in Section 6.2. The results of the empirical analysis are presented in Section 6.3, and finally, Section 6.4 concludes.

## 6.1 The Economic Theory

The money demand function is an important part of the IS-LM (Investment / Saving equilibrium-Liquidity preference / Money supply equilibrium) model. In macroeconomic literature a widely used representation of the long-run money demand relation is as follows.

$$\frac{M}{P} = f(Y, OC) \tag{6.1}$$

$M$ represents the nominal money, $P$ is the price level, $Y$ is the income, and $OC$ is a vector of opportunity cost of holding money. In empirical studies mainly M1, M2 and M3 are used as measures for nominal money. Generally, the consumer price index (CPI) or the GDP deflator are choices for price level. The choice of the variables to constitute the vector of opportunity costs is various. Mainly, nominal short and long-term interest rates are used. In some studies inflation rate is also included as a part of the opportunity costs as it represents the cost of holding money instead of holding assets. Moreover, for countries which are exposed to high inflation, exchange rate can be considered as another opportunity cost variable. In such economies, foreign currency can be a substitute for the national currency.

Following the methodology introduced in Chapter 5, the systems panel cointegration analysis of the money demand function is based on the following model

$$\begin{aligned}\Delta y_{it} &= \alpha_i \beta_i' y_{i,t-1} + \Gamma_{i1}\Delta y_{i,t-1} + \ldots + \Gamma_{i,p_i-1}\Delta y_{i,t-p_i+1} \\ &\quad + C_i d_t + \varepsilon_{it}, \quad i = 1,\ldots,N; \quad t = 1,\ldots,T, \end{aligned} \tag{6.2}$$

in which $y_{it} = (m_{it} - p_{it}, g_{it}, R_{it}, \pi_{it})'$ and $d_t$ contains the deterministic variables. A heterogeneous intercept and a linear trend term are included in the panel VEC model. The trend parameter is included because this will seize the changes in financial technology which influence the money demand, but are neither due to income nor due to opportunity cost of holding money. Note that $m_{it}$ is the logarithm of the nominal M1 level, $p_{it}$ is the logarithm of the CPI, $g_{it}$ is the logarithm of real GDP, $R_{it}$ is either the nominal short-term or the long-term interest rate[1]. The deposit rate defines the short-term interest rate and the government bond yield is used as the long-term interest rate for each country. Finally, $\pi_{it}$ is the annual inflation rate, i.e. $\pi_{it} = 4\Delta p_{it}$ if quarterly data is used. The real variables are produced by dividing the nominal variables by the CPI.

As it will be explained in the next section the panel dataset consists of developing countries along with the European countries. Hence, M1 is used as monetary aggregate because Sriram (2001) indicates that M1 exhibits better behavior than M2 in the empirical studies dealing with developing countries. Moreover, the data corresponding to other definitions of money are not available for all countries under consideration. Note that M3 is preferable when the data consists of just EMU countries as the two-pillar monetary policy strategy of the European Central Bank (ECB) is based on M3.

The long-run money demand relation for the single-equation framework can be formulated as

$$m_{it} - p_{it} = \delta_{0i} + \delta_{1i}t + \beta_1 g_{it} + \beta_2 R_{it} + \beta_3 \pi_{it} + \varepsilon_{it}. \tag{6.3}$$

$\beta_1$, $\beta_2$ and $\beta_3$ parameters denote the income elasticity, the semi-elasticity with respect to the interest rate and the semi-elasticity with respect to the inflation

[1] Throughout the study, the nominal short-term interest rate is represented by $R^s_{it}$ and the nominal long-term interest rate is denoted by $R^l_{it}$.

rate, respectively. According to economic theory, the scale variable $g_{it}$ should have a positive effect on the nominal or real money. Hence, the variables representing the opportunity costs, i.e the nominal interest rate and the inflation rate, should exert a negative effect. Note that the income elasticity is important in determining the rate of monetary expansion consistent with the long-run price stability level. The interest rate semi-elasticity helps to compute the welfare costs of long-run inflation (Mark and Sul, 2003).

## 6.2 The Data

To check for the robustness of the panel SL cointegration test introduced in Chapter 5, different panel datasets with different $T$ and $N$ combinations are evaluated. Throughout this empirical study three different quarterly and seasonally-adjusted panel datasets are used. The first dataset consists of at most 17 countries observed over 1988/Q1-1997/Q4 (i.e. $T = 40$), in the second panel dataset there are at most 18 countries observed over 1990/Q1-1997/Q2 (i.e. $T = 30$), and finally the third panel dataset consists of at most 25 countries observed over 1993/Q1-1997/Q2 (i.e. $T = 18$).

In the tables representing the panel unit root and cointegration tests, the countries included in the panel datasets are coded as follows:

- C11: Austria, Belgium, Finland, France, Germany, Italy, Netherlands, Norway, Portugal, Spain, Switzerland,

- C17 : C11 countries and Australia, Canada, Japan, Korea, New Zealand, US,

- C18 : C17 countries and Denmark,

- C25 : C18 countries and Argentina, Brazil, Indonesia, Malaysia, Mexico, South Africa, Turkey.

The countries are classified mainly according to the availability of the data. C11 countries are extracted from the first and third panel dataset in order to see the robustness of the tests when the cross-section dimension decreases.

The data is provided by the *International Statistical Yearbook* from IMF and OECD databases. A detailed table on the sources of the individual data and definitions of each variable are provided in the Appendix B.2.

## 6.3 Empirical Results

### 6.3.1 Panel Unit Root Tests

Before testing for the existence of a long-run money demand relation, first the integratedness of the variables in the money demand equation should be tested via panel unit root tests. The presence of unit roots in the variables are tested by panel unit root tests of Levin et al. (2002) (henceforth LLC) and Im et al. (2003) (henceforth IPS).

Both tests check under the null hypothesis the existence of a unit root[2] under the assumption that the cross-sections are independent. However, the formulation of the null and alternative hypotheses are different in each test. The LLC test assume under the null hypothesis that there is a common unit root process driving the whole panel. On the contrary, under the null hypothesis of the IPS test, the unit root process is cross-section variant and the rejection of the null hypothesis for IPS test does not necessarily mean that all the cross-sections are stationary. Both tests allow inclusion of deterministic terms into the equation. Even more, the IPS test can not be applied to models without individual-specific deterministic terms.

Basically, both tests are extensions of the ADF statistics to panel data framework. The standardized $t$-bar statistic of IPS is based on the average of the ADF statistics computed for each cross-section separately, and it converges in probability to a standard normal variate if $T \to \infty$ followed by $N \to \infty$. The computation of the LLC test involves three steps. In the first step the ADF statistic is calculated for each cross-section separately and the orthogalized residuals are generated with the help of two auxiliary regressions. In the second step the ratio of the long-run to short-run deviations is estimated e.g with Bartlett kernel. Finally, the test statistic is computed via the orthogalized residuals and the ratio of the long-run to short-run deviations. Their adjusted $t$- statistic is also standard normally distributed as both $T$ and $N$ grow large with $N/T \to 0$. For both, LLC and IPS tests the left tail of the standard normal distribution is used to reject the null hypothesis[3].

In this study, the tests are computed with EViews 6.0. Since both panel unit root tests are based on the ADF regression equation, there is the problem of determining the correct lag order of the process. Throughout this study, the lag order for the unit root processes will be chosen by the Schwarz criterion automatic lag order selection procedure introduced in EViews 6.0.

[2] According to the simulation study of Hlouskova and Wagner (2005) panel stationarity tests perform poorly, therefore they are omitted, from the subsequent analysis.

[3] At the 5% significance level the critical value is -1.65.

Since quarterly data are used, the results are compared to those from selecting a lag order via the Schwarz criterion by restricting the maximum lag order to 4; this results are reported only if proposed lag orders differ. The long-run variance of the regression, which is necessary for the calculation of the semi-parametric panel unit root statistics, is computed via the Newey-West estimator using Bartlett kernel.

The results of the panel unit root tests are presented in Tables 6.1-6.3. Tables include the results for the levels and the first differences of the variables considered. The results based on the dataset from 1988/Q1 to 1997/Q4 consisting of C11 countries indicate that all the variables except the inflation rate are integrated of order one as the nonstationarity of the first differences of the variables is rejected at the 1% significance level (see Table 6.1). Note that only the IPS test indicates the nonstationarity of the inflation rate. When the tests are run for the same time span with C17 countries, all the variables are $I(1)$ (see lower part of Table 6.1).

**Table 6.1:** Panel unit root test results using the quarterly data from 1988/Q1 to 1997/Q4.

| | | Det. | Level | | First Difference | |
|---|---|---|---|---|---|---|
| | Variable | Terms | LLC | IPS | LLC | IPS |
| C11 | *M1* | $c,t$ | 1.48 | 2.39 | -12.73*** | -13.71*** |
| | *GDP* | $c,t$ | -0.79 | 1.63 | -7.39*** | -10.84*** |
| | *log Real M1* | $c,t$ | 1.10 | 2.45 | -14.94*** | -15.00*** |
| | *log Real GDP* | $c,t$ | 0.66 | 0.40 | -13.01*** | -12.97*** |
| | $R^l$ | $c$ | -1.42 | 1.90 | -11.61*** | |
| | $R^s$ | $c$ | 0.71 | 1.96 | -11.43*** | |
| | $\pi$ | $c$ | -1.96** | -2.25** | -24.70*** | |
| | | $c^a$ | 0.46** | -1.27 | | |
| | | None | -3.68*** | | | |
| C17 | *M1* | $c,t$ | 1.04 | 3.55 | -14.70*** | -16.89*** |
| | *GDP* | $c,t$ | -0.17 | 3.16 | -9.52*** | -13.73*** |
| | *log Real M1* | $c,t$ | 1.51 | 3.27 | -20.08*** | -20.53*** |
| | *log Real GDP* | $c,t$ | -1.02 | 0.62 | -16.68*** | -16.90*** |
| | $R^l$ | $c$ | -1.43 | 2.31 | -16.17*** | |
| | $R^s$ | $c$ | 0.83 | 2.26 | -14.76*** | |
| | $\pi$ | $c$ | -4.08 | -5.27 | -29.02*** | |
| | | None | -5.36*** | | | |

*Notes:* [a] *With maximum lag order of 4.*
****, **, * Reject the null hypothesis of nonstationarity at the 1%, 5%, 10% level, respectively.*

The results covering the period from 1990/Q1 to 1997/Q2 differ slightly from the ones above. In Table 6.2 it is apparent that especially with the LLC test at the 10% significance level the nominal and real GDP has no unit root. In general, all the variables except the inflation rate and the short-term interest rate are $I(1)$.

**Table 6.2:** Panel unit root test results using the quarterly data from 1990/Q1 to 1997/Q2 for C18 countries.

| | Det. | Level | | First Difference | |
|---|---|---|---|---|---|
| Variable | Terms | LLC | IPS | LLC | IPS |
| *M1* | $c, t$ | 0.12 | 1.27 | -12.78*** | -14.56*** |
| *GDP* | $c, t$ | -1.88** | -0.86 | -13.03*** | -14.15*** |
| | $c, t^a$ | -1.59* | 0.26 | | |
| *log Real M1* | $c, t$ | -1.36* | -0.10 | -14.22*** | -15.15*** |
| | $c, t^a$ | -0.96 | -0.10 | | |
| *log Real GDP* | $c, t$ | -1.79** | -1.38* | -15.31*** | -16.32*** |
| | $c, t^a$ | -1.62* | -0.74 | | |
| $R^l$ | $c$ | -1.62* | 0.33 | -13.90*** | |
| $R^s$ | $c$ | -0.51 | 1.27 | -13.16*** | |
| $\pi$ | $c$ | -9.10*** | -9.91*** | -28.14*** | |
| | None | -8.16*** | | | |

*Notes:* $^a$ *With maximum lag order of 4.*
****, **, * Reject the null hypothesis of nonstationarity at the 1%, 5%, 10% level, respectively.*

The results in Table 6.3 based on C11 countries with 18 time observations for each cross-section demonstrate that except the log real GDP, the short-term interest rate and the inflation rate, which are stationary at the 1% level, all the variables are integrated of order one. The results with C18 countries do not differ much from the results based on C11 countries. The short-term interest rate and the inflation rate are again $I(0)$. All the other variables are $I(1)$. The results based on the panel dataset consisting of C25 countries indicate that all the variables except the short-term interest rate and the inflation rate have a unit root at least at the 5% significance level. The long-term interest rate is excluded here because the data is not available for all C25 countries. These results based on 25 countries and 18 time observations may be misleading. Hlouskova and Wagner (2005) found out in a simulation study that size and power problems appear if $T$ is too small compared to $N$. In another simulation study, Karlsson and Löthgren (2000) indicated that the LLC test has size distortions for small-$T$ panels. Additionally, they emphasized that the panel unit root tests lose power when a large fraction of the panel is stationary. In order to find out the robustness of the tests, I consider different $T$ and $N$ combinations.

Taking the panel unit root test results into account, in Section 6.3.2, I proceed with testing for panel cointegration knowing that M1, GDP, $R^s$, $R^l$ are all $I(1)$.

**Table 6.3:** Panel unit root test results using the quarterly data from 1993/Q1 to 1997/Q2.

| | Variable | Det. Terms | Level LLC | Level IPS | First Difference LLC | First Difference IPS |
|---|---|---|---|---|---|---|
| C11 | *M1* | $c, t$ | 1.78 | 1.83 | -8.02*** | -8.06*** |
| | *GDP* | $c, t$ | -0.03 | -0.69 | -12.07*** | -11.47*** |
| | *log Real M1* | $c, t$ | 0.69 | 0.73 | -7.77*** | -7.97*** |
| | *log Real GDP* | $c, t$ | -3.91*** | -3.73*** | -8.91*** | -8.43*** |
| | $R^s$ | $c$ | -9.76*** | -5.42*** | -8.19*** | |
| | $\pi$ | $c$ | -2.78*** | -2.05** | -9.17*** | |
| | | None | -3.06*** | | | |
| C18 | *M1* | $c, t$ | -0.66 | 0.31 | -10.56*** | -9.11*** |
| | *GDP* | $c$ | -1.05 | -0.12 | -18.48*** | -14.12*** |
| | *log Real M1* | $c, t$ | -0.81 | -1.32* | -11.46*** | -9.93*** |
| | *log Real GDP* | $c, t$ | -2.38*** | -2.05** | -12.69*** | -12.10*** |
| | | $c, t^a$ | -1.59* | -0.98 | | |
| | $R^s$ | $c$ | -6.07*** | -2.61*** | -10.01*** | |
| | $\pi$ | $c$ | -7.60*** | -8.67*** | -22.04*** | |
| | | None | -6.08*** | | | |
| C25 | *M1* | $c, t$ | 0.46 | 1.34 | -10.70*** | -9.09*** |
| | *GDP* | $c, t$ | 0.88 | 1.11 | -15.38*** | -13.53*** |
| | *log Real M1* | $c, t$ | -0.88* | -1.57* | -13.13*** | -11.45*** |
| | *log Real GDP* | $c, t$ | -2.94*** | -2.28** | -14.62*** | -13.23*** |
| | | $c, t^a$ | -2.28** | -1.28 | | |
| | $R^s$ | $c$ | -6.24*** | -2.99*** | -12.27*** | |
| | $\pi$ | $c$ | -8.40*** | -9.93*** | -24.97*** | |
| | | None | -6.41*** | | | |

*Notes:* [a] *With maximum lag order of 4.*
****, **, * Reject the null hypothesis of nonstationarity at the 1%, 5%, 10% level, respectively.*

### 6.3.2 Panel Cointegration Tests

To test the existence of a stable long-run money demand relation, in addition to the maximum-likelihood-based panel SL test proposed in Chapter 5, I apply also the residual-based tests of Pedroni (1999), which were introduced in Chapter 2. The LLL test is excluded from the subsequent analysis because the finite sample properties of the LLL test are not better than the properties of the panel SL test. The analysis of the panel SL test is based on Equation (6.2) and the analysis of the residual-based tests of Pedroni (1999) are based on Equation (6.3). Among the seven panel cointegration statistics suggested by Pedroni (1999), I choose the panel-$\rho$, panel-$t$, group-$\rho$ and group-$t$ statistics as their finite sample properties were analyzed in Chapter 3. All these test statistics mentioned above have a limiting standard normal distribution after the appropriate standardization with the asymptotic moments. Thus, the four test statistics of Pedroni (1999) converge to negative infinity under the alternative hypothesis.

The main difference between the panel SL and residual-based tests of Pedroni (1999) is that with the panel SL test, I can decide how many cointegrating relations there are among different variables.

The results of the panel cointegration tests are listed in Tables 6.4-6.13. Please note that for each cointegration test there is an intercept and a linear time trend in both the heterogeneous VAR and panel regression models. The maximum lag order of the individual VAR models (Equation (5.1) and (5.2)) and the panel ADF regression (Section 2.1.1) for the panel-$t$ and group-$t$ statistics of Pedroni (1999) are 4 as the analysis is based on quarterly data. The appropriate lag order is automatically determined for the tests of Pedroni by EViews 6.0 with the Schwarz criterion. The bandwidth selection for the semi-parametric tests of Pedroni (Equation (2.3)) is performed using the Bartlett kernel. Hence, the appropriate lag order of the individual VAR models are determined by the Schwarz criterion using JMulTi v4.21 (see Lütkepohl and Krätzig, 2004) because it has a user friendly feature to apply the cointegration tests of Saikkonen and Lütkepohl (2000a).

The results in Tables 6.4-6.7 are based on the period from 1988:Q1 to 1997:Q4, and four different representations of the money demand relations are analyzed. The findings in Table 6.4 including the long-run interest rate as the only opportunity cost variable reveal that 9 out of 17 OECD countries do not have a long-run stable money demand relation. Hence, Australia, Belgium, Finland, Germany, Portugal and US have one cointegrating relation. The panel cointegration tests have conflicting outcomes. The panel SL test proposed in Chapter 5 finds one cointegrating relation at the 1% significance level. This is an evidence for the existence of the long-run stationary relation between $m_{it} - p_{it}$, $y_{it}$ and $R^l_{it}$. However, the existence of a stationary cointegrating relation cannot be justified using the residual-based tests of Pedroni (1999). Only the parametric panel-$t$ test rejects the null hypothesis of no cointegration at most at the 5% significance level. Note that a rejection of the panel test statistics, i.e. panel-$t$ and panel-$\rho$, means that there is a common cointegrating relation for the whole panel data, whereas a rejection of the group test statistics, i.e. group-$t$ and group-$\rho$, indicates that not all the cross-sections belonging to the panel data have a heterogeneous cointegrating relation.

**Table 6.4:** Cointegration test results for the data from 1988/Q1 to 1997/Q4 when the long-term interest rate is the opportunity cost.

| Country by Country Trace Test Statistics | | | | | |
|---|---|---|---|---|---|
| Country | lag | $r=0$ | $r=1$ | $r=2$ | rank |
| Australia | 1 | 32.71 | 14.28 | 0.36 | 1 |
| Austria | 1 | 23.21 | 5.75 | 0.09 | 0 |
| Belgium | 1 | 34.19 | 8.09 | 1.99 | 1 |
| Canada | 1 | 24.90 | 3.74 | 0.56 | 0 |
| Finland | 2 | 28.60 | 9.84 | 0.03 | 1 |
| France | 1 | 18.63 | 2.58 | 2.20 | 0 |
| Germany | 1 | 33.78 | 11.80 | 1.10 | 1 |
| Italy | 1 | 16.24 | 5.55 | 0.62 | 0 |
| Japan | 1 | 19.72 | 10.46 | 1.41 | 0 |
| Korea | 1 | 13.28 | 7.58 | 2.35 | 0 |
| Netherlands | 1 | 20.82 | 4.89 | 2.50 | 0 |
| New Zealand | 1 | 36.04 | 16.44 | 2.44 | 2 |
| Norway | 2 | 28.44 | 3.60 | 0.03 | 0 |
| Portugal | 1 | 29.99 | 9.99 | 1.57 | 1 |
| Spain | 2 | 13.96 | 7.44 | 0.74 | 0 |
| Switzerland | 1 | 46.01 | 17.53 | 0.01 | 2 |
| US | 1 | 45.43 | 10.17 | 6.89 | 1 |
| Panel Tests | | $r=0$ | $r=1$ | $r=2$ | |
| C11 | | 4.84 | -9.91 | -2.73 | 1*** |
| C17 | | 6.56 | -0.14 | -2.93 | 1*** |
| Pedroni Tests | | panel-$\rho$ | panel-$t$ | group-$\rho$ | group-$t$ |
| C11 | | 0.36 | -1.70** | 1.35 | -1.08 |
| C17 | | 0.76 | -1.62* | 1.56 | -0.96 |

*Notes: ***, **, * Reject the null hypothesis at the 1%, 5%, 10% level. For country by country trace tests the critical values at 5% significance level are 28.47, 15.92, 6.83 to test the hypotheses $r=0$, $r=1$ and $r=2$, respectively. Critical values at the 5% significance level for the panel SL test and the tests of Pedroni are 1.645 and -1.645, respectively.*

From Table 6.5 one can draw similar conclusions whenever the short-term interest rate is the opportunity cost measure, i.e. the panel SL test rejects the null hypothesis of no cointegration at the 1% level, and only the panel-$t$ statistic rejects the null hypothesis of no cointegration at the 10% level.

Tables 6.6 and 6.7 list the outcomes of the money demand models with inflation rate as a second opportunity cost measure along with the short-or long-term interest rates. For country by country maximum-likelihood-based tests 8 countries indicate the existence of 2 cointegrating relations at the 5% level. One of these cointegrating relations is the long-run money demand relation, and the second relation can be the long-run stationary relation between $R^l_{it}$ and $\pi_{it}$, which is the Fisher relation analyzed already in Chapter 3. Either with the long-term interest rate or the short-term interest rate as the first opportunity cost variable the panel SL test points out the existence of two cointegrating relations. Hence, the tests of Pedroni do not show again evidence for the existence of a long-run stationary relation.

Now I reduce the number of time observations for each country to 30 and apply the tests on a panel data consisting of C18 OECD countries over the period from 1990/Q1 to 1997/Q2. Individual trace tests in Tables 6.8 and 6.9 with long-term or short-term interest rates as the opportunity cost variables indicate that half of the countries do not have a long-run money demand relation. Hence, almost the entire other half of the countries show evidence for an existing cointegrating relation. At the 5% level, Denmark, Finland, Spain and the US have one cointegrating relation. Using $R^l$ as the opportunity cost variable Germany has one stationary long-run relation, but with $R^s$ as the opportunity cost variable, there are two cointegrating relations. The first cointegrating relation is the long-run money demand while the second relation can represent the stationarity of $R^s$. In both tables the panel SL test indicates the existence of the long-run money demand relation, whereas the existence of one cointegrating relation is approved only with the panel-$t$ and group-$t$ tests at the 1% significance level.

After the inclusion of the inflation rate as a second opportunity cost variable in the system, almost 9 out of C18 OECD countries have one cointegrating relation. If the long-term interest rate is the first opportunity cost variable the panel SL test finds two cointegrating relations at the 5% level (see Table 6.10). However, if the short-term interest rate is the first opportunity cost measure, the existence of two cointegrating relations is rejected at the 1% significance level (see Table 6.11). Thus, the second and third cointegrating relations may be the stationarity of the short-term interest and/or inflation rates as also depicted by the panel unit root test results in Section 6.3.1. With the inclusion of an additional opportunity cost variable, Pedroni's tests find more evidence for the existence of a cointegrating relation. In Tables 6.10 and 6.11 the panel-$t$ test rejects the null hypothesis of no cointegration at the 1% significance level, whereas the group-$t$ test rejects it only at the 5% significance level. However, these statistics do not emphasize how many cointegrating relations there are among the variables.

**Table 6.5:** Cointegration test results for the data from 1988/Q1 to 1997/Q4 when the short-term interest rate is the opportunity cost.

| Country by Country Trace Test Statistics | | | | | |
|---|---|---|---|---|---|
| Country | lag | $r = 0$ | $r = 1$ | $r = 2$ | rank |
| Australia | 1 | 42.83 | 14.61 | 0.39 | 1 |
| Austria | 1 | 39.01 | 7.27 | 0.01 | 1 |
| Belgium | 1 | 21.14 | 8.31 | 0.83 | 0 |
| Canada | 1 | 28.48 | 4.38 | 2.33 | 1 |
| Finland | 2 | 31.56 | 6.35 | 0.20 | 1 |
| France | 1 | 12.58 | 6.89 | 1.35 | 0 |
| Germany | 3 | 29.10 | 14.18 | 0.04 | 1 |
| Italy | 1 | 14.08 | 6.78 | 0.61 | 0 |
| Japan | 1 | 18.19 | 6.90 | 0.74 | 0 |
| Korea | 1 | 22.69 | 5.94 | 0.94 | 0 |
| Netherlands | 1 | 13.45 | 4.68 | 0.87 | 0 |
| New Zealand | 2 | 38.37 | 0.65 | 0.09 | 1 |
| Norway | 2 | 15.11 | 5.62 | 0.53 | 0 |
| Portugal | 1 | 39.46 | 4.19 | 0.86 | 1 |
| Spain | 2 | 16.06 | 7.40 | 0.63 | 0 |
| Switzerland | 1 | 50.39 | 23.26 | 0.13 | 2 |
| US | 1 | 36.36 | 6.08 | 0.16 | 1 |
| Panel Tests | | $r = 0$ | $r = 1$ | $r = 2$ | |
| C11 | | 4.17 | -0.27 | -3.43 | 1*** |
| C17 | | 6.70 | -1.20 | -4.10 | 1*** |
| Pedroni Tests | | panel-$\rho$ | panel-$t$ | group-$\rho$ | group-$t$ |
| C11 | | 0.37 | -1.11 | 1.27 | -0.80 |
| C17 | | 0.45 | -1.50* | 1.37 | -0.87 |

*Notes: ***, **, * Reject the null hypothesis at the 1%, 5%, 10% level.*

The existence of a stable money demand is also tested using 25 OECD countries with 18 observations from 1993/Q2 to 1997/Q2. Only the short-term interest rate and the inflation rate can be used as opportunity cost variables because the data on the long-term interest rates are not available for some developing countries (see Tables 6.12-6.13). For the residual-based tests the maximum lag order is still 4. However, the maximum possible lag order of the individual VAR models is limited to 2 because of the number of observations. Considering the individual trace statistics, in Table 6.12 there is no evidence for a cointegrating relation for more than the half of the countries, and the other countries have one cointegrating relation. The panel SL test also provides evidence for the existence of one cointegrating relation.

**Table 6.6:** Cointegration test results for the data from 1988/Q1 to 1997/Q4 when the long-term interest rate and the inflation rate are the opportunity costs.

| Country by Country Trace Test Statistics | | | | | | |
|---|---|---|---|---|---|---|
| Country | lag | $r = 0$ | $r = 1$ | $r = 2$ | $r = 3$ | rank |
| Australia | 1 | 63.78 | 29.71 | 16.55 | 0.55 | 3 |
| Austria | 3 | 71.50 | 30.45 | 5.85 | 0.58 | 2 |
| Belgium | 1 | 58.76 | 34.95 | 8.23 | 2.02 | 2 |
| Canada | 1 | 61.74 | 29.55 | 4.23 | 0.42 | 2 |
| Finland | 2 | 70.14 | 28.79 | 10.34 | 0.03 | 2 |
| France | 1 | 50.20 | 18.72 | 4.25 | 2.17 | 1 |
| Germany | 1 | 73.96 | 35.80 | 12.63 | 1.10 | 2 |
| Italy | 1 | 38.70 | 17.30 | 8.02 | 0.41 | 0 |
| Japan | 1 | 65.18 | 19.75 | 8.50 | 2.22 | 1 |
| Korea | 1 | 39.20 | 15.48 | 7.42 | 1.77 | 0 |
| Netherlands | 1 | 44.50 | 20.55 | 6.49 | 2.40 | 0 |
| New Zealand | 1 | 61.98 | 31.77 | 13.96 | 2.41 | 2 |
| Norway | 1 | 38.23 | 16.52 | 5.40 | 0.87 | 0 |
| Portugal | 1 | 63.47 | 28.99 | 10.17 | 2.05 | 2 |
| Spain | 4 | 63.65 | 21.26 | 9.91 | 3.67 | 1 |
| Switzerland | 3 | 85.11 | 42.73 | 17.81 | 2.26 | 3 |
| US | 1 | 110.39 | 44.69 | 12.85 | 5.64 | 2 |
| Panel Tests | | $r = 0$ | $r = 1$ | $r = 2$ | $r = 3$ | |
| C11 | | 6.56 | 4.97 | 0.07 | -1.68 | 2*** |
| C17 | | 9.99 | 6.61 | 0.71 | -1.79 | 2*** |
| Pedroni Tests | | panel-$\rho$ | panel-$t$ | group-$\rho$ | group-$t$ | |
| C11 | | 1.61 | 0.37 | 2.54 | 0.21 | |
| C17 | | 2.18 | 1.25 | 3.11 | 1.25 | |

*Notes: ***, **, * Reject the null hypothesis at the 1%, 5%, 10% level. For country by country trace tests the critical values at 5% significance level are 45.13, 28.47, 15.92, 6.83 to test the hypotheses $r = 0$, $r = 1$, $r = 2$ and $r = 3$, respectively.*

Adding inflation rate as the second opportunity cost variable the panel SL test rejects the null hypothesis of one cointegrating relation at the 1% significance level (see Table 6.13). Moreover, the results are also valid for the panel datasets consisting of C11 and C18, which are presented in Tables 6.12 and 6.13. Furthermore, with the increase in $N$, the panel-$t$ and group-$t$ statistics reject the null hypothesis of no cointegration.

Note that the panel-$\rho$ and group-$\rho$ tests cannot find evidence for a cointegrating relation because they have low power for small $T$ even when $N$ increases (Please refer to the simulation results in Chapter 3.).

In the following section the long-run money demand equation is estimated with the panel DOLS estimation method developed by Mark and Sul (2003). After the explanation of the estimation method, the estimated income elasticities, interest rate semi-elasticities and the inflation rate semi-elasticities will be presented.

**Table 6.7:** Cointegration test results for the data from 1988/Q1 to 1997/Q4 when the short-term interest rate and the inflation rate are the opportunity costs.

| Country by Country Trace Test Statistics | | | | | | |
|---|---|---|---|---|---|---|
| Country | lag | $r = 0$ | $r{=}1$ | $r = 2$ | $r = 3$ | rank |
| Australia | 1 | 69.09 | 28.78 | 15.13 | 0.40 | 2 |
| Austria | 1 | 125.41 | 34.87 | 6.12 | 0.01 | 2 |
| Belgium | 1 | 38.22 | 21.36 | 7.82 | 0.90 | 0 |
| Canada | 1 | 61.53 | 27.40 | 4.33 | 2.67 | 1 |
| Finland | 2 | 66.73 | 28.75 | 7.29 | 0.19 | 2 |
| France | 1 | 51.25 | 13.46 | 6.80 | 1.71 | 1 |
| Germany | 1 | 79.64 | 35.27 | 3.03 | 0.02 | 2 |
| Italy | 1 | 37.59 | 15.23 | 7.13 | 0.75 | 0 |
| Japan | 1 | 63.64 | 18.47 | 4.64 | 1.46 | 1 |
| Korea | 1 | 48.30 | 27.08 | 5.68 | 0.95 | 1 |
| Netherlands | 1 | 39.40 | 15.80 | 5.62 | 0.83 | 0 |
| New Zealand | 2 | 73.98 | 36.04 | 2.13 | 0.15 | 2 |
| Norway | 1 | 33.83 | 10.06 | 4.50 | 2.24 | 0 |
| Portugal | 1 | 75.96 | 40.90 | 4.29 | 1.38 | 2 |
| Spain | 4 | 41.11 | 16.05 | 7.74 | 0.33 | 0 |
| Switzerland | 3 | 72.61 | 37.73 | 6.80 | 0.39 | 2 |
| US | 1 | 92.24 | 36.87 | 7.20 | 0.15 | 2 |
| Panel Tests | | $r = 0$ | $r = 1$ | $r = 2$ | $r = 3$ | |
| C11 | | 6.75 | 3.46 | -2.53 | -3.04 | $2^{***}$ |
| C17 | | 10.42 | 5.57 | -2.98 | -3.66 | $2^{***}$ |
| Pedroni Tests | | panel-$\rho$ | panel-$t$ | group-$\rho$ | group-$t$ | |
| C11 | | 1.63 | 0.19 | 2.52 | 0.54 | |
| C17 | | 1.82 | 0.60 | 2.79 | 1.18 | |

*Notes: ***, **, * Reject the null hypothesis at the 1%, 5%, 10% level.*

### 6.3.3 Estimation of Long-Run Money Demand

In this section Equation (6.3) is estimated with the method introduced in Mark and Sul (2003). In their study Mark and Sul (2003) proposed a panel DOLS estimator for a homogeneous cointegrating vector, based on a balanced panel. The estimator is totally parametric and easy to implement. Since Kao and Chiang (2000) concluded that the panel DOLS estimator outperforms both panel OLS and panel FMOLS estimators, the panel DOLS estimator of Mark and Sul (2003) is preferred against the panel FMOLS estimator of Pedroni (2000).

**Table 6.8:** Cointegration test results for the data from 1990/Q1 to 1997/Q2 when the long-term interest rate is the opportunity cost.

| Country by Country Trace Test Statistics | | | | | |
|---|---|---|---|---|---|
| Country | lag | $r = 0$ | $r = 1$ | $r = 2$ | rank |
| Australia | 4 | 24.38 | 4.76 | 0.03 | 0 |
| Austria | 1 | 29.47 | 7.16 | 2.98 | 1 |
| Belgium | 1 | 20.88 | 16.28 | 3.44 | 0 |
| Canada | 1 | 23.61 | 5.23 | 0.14 | 0 |
| Denmark | 2 | 38.46 | 13.29 | 0.56 | 1 |
| Finland | 4 | 32.73 | 6.73 | 0.01 | 1 |
| France | 1 | 12.64 | 3.65 | 0.68 | 0 |
| Germany | 4 | 31.21 | 11.58 | 0.09 | 1 |
| Italy | 1 | 14.96 | 3.64 | 1.61 | 0 |
| Japan | 1 | 16.62 | 5.85 | 0.16 | 0 |
| Korea | 1 | 17.38 | 3.87 | 1.72 | 0 |
| Netherlands | 1 | 15.55 | 3.75 | 1.08 | 0 |
| New Zealand | 1 | 35.20 | 13.87 | 0.26 | 1 |
| Norway | 4 | 32.33 | 7.09 | 0.97 | 1 |
| Portugal | 1 | 21.08 | 5.06 | 1.74 | 0 |
| Spain | 4 | 41.81 | 12.06 | 0.63 | 1 |
| Switzerland | 2 | 39.72 | 12.44 | 0.32 | 1 |
| US | 1 | 41.39 | 12.06 | 5.64 | 1 |
| Panel Tests | | $r = 0$ | $r = 1$ | $r = 2$ | |
| C18 | | 6.58 | -0.79 | -3.01 | 1*** |
| Pedroni Tests | | panel-$\rho$ | panel-$t$ | group-$\rho$ | group-$t$ |
| C18 | | 0.49 | -3.91*** | 1.48 | -2.58*** |

*Notes: ***, **, * Reject the null hypothesis at the 1%, 5%, 10% level.*

**Table 6.9:** Cointegration test results for the data from 1990/Q1 to 1997/Q2 when the short-term interest rate is the opportunity cost.

| Country by Country Trace Test Statistics | | | | | |
|---|---|---|---|---|---|
| Country | lag | $r = 0$ | $r = 1$ | $r = 2$ | rank |
| Australia | 4 | 31.62 | 11.06 | 1.72 | 1 |
| Austria | 1 | 17.39 | 4.20 | 0.76 | 0 |
| Belgium | 1 | 14.13 | 10.34 | 2.27 | 0 |
| Canada | 1 | 24.04 | 5.60 | 0.01 | 0 |
| Denmark | 4 | 33.26 | 13.75 | 4.98 | 1 |
| Finland | 1 | 38.75 | 6.65 | 2.62 | 1 |
| France | 1 | 15.91 | 3.98 | 0.65 | 0 |
| Germany | 4 | 44.88 | 20.15 | 1.90 | 2 |
| Italy | 1 | 13.99 | 3.11 | 1.96 | 0 |
| Japan | 1 | 13.24 | 7.35 | 2.64 | 0 |
| Korea | 1 | 18.74 | 4.24 | 1.11 | 0 |
| Netherlands | 4 | 31.78 | 12.67 | 1.43 | 1 |
| New Zealand | 2 | 43.94 | 19.11 | 0.33 | 2 |
| Norway | 1 | 16.99 | 4.45 | 1.71 | 0 |
| Portugal | 1 | 33.37 | 2.67 | 1.08 | 1 |
| Spain | 4 | 40.44 | 10.62 | 0.55 | 1 |
| Switzerland | 1 | 26.56 | 10.06 | 2.18 | 0 |
| US | 4 | 50.18 | 9.48 | 3.36 | 1 |
| Panel Tests | | $r = 0$ | $r = 1$ | $r = 2$ | |
| C18 | | 7.46 | -0.08 | -1.96 | 1*** |
| Pedroni Tests | | panel-$\rho$ | panel-$t$ | group-$\rho$ | group-$t$ |
| C18 | | -0.37 | -3.40*** | 0.65 | -2.40*** |

*Notes: ***, **, * Reject the null hypothesis at the 1%, 5%, 10% level.*

Mark and Sul (2003) extended the DOLS method of Saikkonen (1991) and Stock and Watson (1993) to estimate the cointegrating vector in panel data. They based their estimation procedure on the following triangular representation.

$$y_{it} = \delta_{0i} + \delta_{1i}t + x'_{it}\beta + \varepsilon_{it}, \quad (6.4)$$

$$x_{it} = x_{i,t-1} + \nu_{it}, \quad (6.5)$$

in which $x_{it}$ is a $K$-dimensional vector of regressors. Note that for the money demand relation $y_{it} = m_{it} - p_{it}$ and $x_{it} = (g_{it}, R_{it}, \pi_{it})$ is the vector of regressors.

**Table 6.10:** Cointegration test results for the data from 1990/Q1 to 1997/Q2 when the long-term interest rate and the inflation rate are the opportunity costs.

| Country by Country Trace Test Statistics | | | | | | |
|---|---|---|---|---|---|---|
| Country | lag | $r=0$ | $r=1$ | $r=2$ | $r=3$ | rank |
| Australia | 4 | 71.50 | 31.45 | 8.62 | 5.02 | 2 |
| Austria | 4 | 51.07 | 37.75 | 11.36 | 1.18 | 2 |
| Belgium | 4 | 55.48 | 19.09 | 12.27 | 2.39 | 1 |
| Canada | 1 | 56.65 | 18.10 | 4.61 | 0.12 | 1 |
| Denmark | 4 | 64.26 | 25.98 | 10.58 | 0.12 | 1 |
| Finland | 4 | 61.10 | 19.21 | 14.07 | 0.06 | 1 |
| France | 4 | 85.50 | 23.87 | 6.35 | 0.67 | 1 |
| Germany | 4 | 54.07 | 26.92 | 11.28 | 0.21 | 1 |
| Italy | 4 | 61.92 | 33.96 | 17.40 | 0.67 | 3 |
| Japan | 1 | 32.11 | 16.83 | 8.30 | 0.66 | 0 |
| Korea | 1 | 33.60 | 17.29 | 3.76 | 1.66 | 0 |
| Netherlands | 1 | 35.96 | 17.00 | 4.66 | 1.31 | 0 |
| New Zealand | 1 | 57.04 | 26.32 | 8.29 | 1.01 | 1 |
| Norway | 4 | 71.85 | 46.64 | 8.69 | 0.00 | 2 |
| Portugal | 4 | 73.72 | 23.08 | 8.13 | 1.91 | 1 |
| Spain | 4 | 72.36 | 59.44 | 25.49 | 1.04 | 3 |
| Switzerland | 1 | 49.70 | 20.31 | 15.01 | 1.35 | 1 |
| US | 4 | 84.64 | 39.79 | 5.82 | 0.09 | 2 |
| Panel Tests | | $r=0$ | $r=1$ | $r=2$ | $r=3$ | |
| C18 | | 8.21 | 7.18 | 1.53 | -3.30 | 2** |
| Pedroni Tests | | panel-$\rho$ | panel-$t$ | group-$\rho$ | group-$t$ | |
| C18 | | 1.74 | -2.74*** | 2.78 | -1.76** | |

*Notes: ***, **, * Reject the null hypothesis at the 1%, 5%, 10% level.*

The main assumption of the panel DOLS approach of Mark and Sul (2003) is the homogeneity of the cointegrating vector over individuals. Thus, the intercept and the trend terms are allowed to vary over cross-sections. The idiosyncratic error terms are cross-sectionally independent, but they may be serially correlated. Moreover, $\mathrm{w}_{it} = (\varepsilon_{it}, \nu'_{it})'$ is independent across $N$ individuals, and has a moving average representation, in which the moving average parameters may differ across $N$ individuals. This includes additional heterogeneity into the system.

To dispose the endogeneity problem, which is the due to the correlation of $\varepsilon_{it}$ with $p_i$ leads and lags of $\nu_{it}$, $\varepsilon_{it}$ is projected onto the $p_i$ leads and lags of $\nu_{it} = \Delta x_{it}$ for individual $i$.

**Table 6.11:** Cointegration test results for the data from 1990/Q1 to 1997/Q2 when the short-term interest rate and the inflation rate are the opportunity costs.

| Country by Country Trace Test Statistics | | | | | | |
|---|---|---|---|---|---|---|
| Country | lag | $r=0$ | $r=1$ | $r=2$ | $r=3$ | rank |
| Australia | 4 | 83.16 | 46.57 | 11.25 | 1.94 | 2 |
| Austria | 4 | 66.44 | 26.81 | 12.99 | 2.98 | 1 |
| Belgium | 1 | 59.27 | 13.53 | 9.93 | 2.62 | 1 |
| Canada | 1 | 55.22 | 17.94 | 3.60 | 0.00 | 1 |
| Denmark | 4 | 69.80 | 24.62 | 15.80 | 0.07 | 1 |
| Finland | 4 | 43.98 | 24.11 | 15.75 | 1.20 | 0 |
| France | 1 | 54.20 | 17.45 | 4.43 | 1.00 | 1 |
| Germany | 4 | 77.05 | 29.86 | 27.17 | 3.28 | 3 |
| Italy | 4 | 84.50 | 25.98 | 13.41 | 0.19 | 1 |
| Japan | 1 | 30.31 | 12.55 | 7.38 | 3.79 | 0 |
| Korea | 1 | 37.91 | 28.86 | 4.27 | 1.32 | 0 |
| Netherlands | 1 | 37.81 | 20.72 | 8.65 | 0.76 | 0 |
| New Zealand | 4 | 79.43 | 14.84 | 10.67 | 0.00 | 1 |
| Norway | 4 | 60.80 | 28.59 | 14.29 | 0.25 | 2 |
| Portugal | 1 | 53.60 | 21.60 | 2.19 | 1.50 | 1 |
| Spain | 4 | 79.58 | 58.42 | 15.31 | 1.70 | 2 |
| Switzerland | 2 | 42.27 | 17.88 | 7.73 | 2.18 | 0 |
| US | 4 | 62.50 | 27.72 | 16.79 | 1.21 | 3 |
| Panel Tests | | $r=0$ | $r=1$ | $r=2$ | $r=3$ | |
| C18 | | 8.43 | 5.19 | 2.59 | -2.67 | 3*** |
| Pedroni Tests | | panel-$\rho$ | panel-$t$ | group-$\rho$ | group-$t$ | |
| C18 | | 1.11 | -3.01*** | 2.05 | -2.11** | |

*Notes: ***, **, * Reject the null hypothesis at the 1%, 5%, 10% level.*

$$\varepsilon_{it} = \sum_{s=-p_i}^{p_i} \Delta x'_{i,t-s}\lambda_{is} + u_{it} = z'_{it}\lambda_i + u_{it}. \qquad (6.6)$$

$\lambda_{is}$ is a $K$-dimensional vector of projection coefficients, $\lambda_i = (\lambda'_{i,-p_i}, \dots, \lambda'_{i0}, \dots, \lambda'_{i,p_i})'$ is a $2(p_i+1)K$-dimensional vector and $z_{it} = (\Delta x'_{i,t-p_i}, \dots, \Delta x'_{it}, \dots, \Delta x'_{i,t+p_i})'$ is a $2(p_i+1)K$-dimensional vector of leads and lags of $\Delta x_{it}$. In the next step, (6.6) is inserted into (6.4) which leads to

$$y_{it} = \delta_{0i} + \delta_{1i}t + x'_{it}\beta + z'_{it}\lambda_i + u_{it}. \qquad (6.7)$$

**Table 6.12:** Cointegration test results for the data from 1993/Q1 to 1997/Q2 when the short-term interest rate is the opportunity cost. The maximum lag order of the VAR model is 2.

| Country by Country Trace Test Statistics | | | | | |
|---|---|---|---|---|---|
| Country | lag | $r = 0$ | $r = 1$ | $r = 2$ | rank |
| Australia | 1 | 37.26 | 4.19 | 0.03 | 1 |
| Austria | 1 | 20.23 | 9.17 | 4.90 | 0 |
| Belgium | 2 | 25.14 | 7.21 | 3.17 | 0 |
| Canada | 2 | 29.18 | 9.27 | 0.25 | 1 |
| Denmark | 2 | 30.62 | 4.93 | 0.27 | 1 |
| Finland | 2 | 21.00 | 12.02 | 0.34 | 0 |
| France | 2 | 36.37 | 4.20 | 0.55 | 1 |
| Germany | 2 | 20.67 | 6.85 | 0.69 | 0 |
| Italy | 2 | 20.09 | 19.20 | 0.49 | 0 |
| Japan | 2 | 41.69 | 12.30 | 0.38 | 1 |
| Korea | 1 | 27.68 | 7.82 | 0.77 | 0 |
| Netherlands | 1 | 23.86 | 4.63 | 1.01 | 0 |
| New Zealand | 2 | 28.95 | 1.74 | 2.67 | 1 |
| Norway | 1 | 18.71 | 11.28 | 1.95 | 0 |
| Portugal | 1 | 35.72 | 3.76 | 1.48 | 1 |
| Spain | 2 | 23.93 | 8.10 | 3.90 | 0 |
| Switzerland | 2 | 31.04 | 13.24 | 0.10 | 1 |
| US | 1 | 40.69 | 2.37 | 0.12 | 1 |
| Argentina | 2 | 23.89 | 6.00 | 3.49 | 0 |
| Brazil | 1 | 17.38 | 12.39 | 6.56 | 0 |
| Indonesia | 2 | 31.56 | 6.27 | 3.92 | 1 |
| Malaysia | 2 | 27.38 | 5.76 | 0.30 | 0 |
| Mexico | 1 | 27.04 | 13.41 | 0.89 | 0 |
| South Africa | 1 | 33.80 | 1.63 | 0.98 | 1 |
| Turkey | 2 | 24.63 | 12.27 | 0.91 | 0 |
| Panel Tests | | $r = 0$ | $r = 1$ | $r = 2$ | |
| C11 | | 3.79 | 0.12 | -1.61 | 1*** |
| C18 | | 7.62 | -1.72 | -2.89 | 1*** |
| C25 | | 8.46 | -1.25 | -2.63 | 1*** |
| Pedroni Tests | | panel-$\rho$ | panel-$t$ | group-$\rho$ | group-$t$ |
| C11 | | 0.83 | -1.27 | 2.01 | -1.82** |
| C18 | | 0.55 | -4.44*** | 2.18 | -4.46*** |
| C25 | | 1.82 | -2.20** | 2.44 | -6.74*** |

*Notes: ***, **, * Reject the null hypothesis at the 1%, 5%, 10% level.*

**Table 6.13:** Cointegration test results for the data from 1993/Q1 to 1997/Q2 when the short-term interest rate and the inflation rate are the opportunity costs. The maximum lag order of the VAR model is 2.

| | | Country by Country Trace Test Statistics | | | | |
|---|---|---|---|---|---|---|
| Country | lag | $r=0$ | $r=1$ | $r=2$ | $r=3$ | rank |
| Australia | 2 | 78.22 | 26.20 | 16.51 | 1.34 | 1 |
| Austria | 1 | 53.70 | 19.92 | 8.06 | 1.44 | 1 |
| Belgium | 2 | 44.14 | 21.54 | 7.78 | 4.53 | 0 |
| Canada | 2 | 49.64 | 30.32 | 8.89 | 0.23 | 2 |
| Denmark | 2 | 56.84 | 24.64 | 8.66 | 0.39 | 1 |
| Finland | 2 | 40.33 | 29.86 | 6.82 | 0.02 | 0 |
| France | 2 | 56.91 | 24.80 | 11.06 | 0.80 | 1 |
| Germany | 1 | 41.27 | 22.29 | 4.15 | 1.08 | 0 |
| Italy | 2 | 59.85 | 33.74 | 9.40 | 0.63 | 2 |
| Japan | 2 | 74.20 | 21.42 | 10.77 | 1.42 | 1 |
| Korea | 1 | 44.05 | 23.76 | 6.74 | 0.71 | 0 |
| Netherlands | 1 | 37.46 | 20.52 | 7.35 | 0.50 | 0 |
| New Zealand | 2 | 49.46 | 28.57 | 8.50 | 0.02 | 2 |
| Norway | 2 | 53.04 | 15.67 | 5.47 | 0.15 | 1 |
| Portugal | 2 | 78.27 | 36.08 | 5.55 | 2.00 | 2 |
| Spain | 2 | 46.96 | 36.93 | 17.34 | 6.56 | 3 |
| Switzerland | 2 | 67.37 | 24.42 | 13.04 | 1.20 | 1 |
| US | 2 | 78.58 | 28.16 | 14.38 | 1.93 | 1 |
| Argentina | 2 | 51.31 | 16.56 | 4.67 | 1.27 | 1 |
| Brazil | 2 | 65.17 | 31.26 | 6.46 | 3.18 | 2 |
| Indonesia | 2 | 61.77 | 28.34 | 6.37 | 0.16 | 1 |
| Malaysia | 2 | 49.56 | 25.95 | 4.30 | 0.53 | 1 |
| Mexico | 2 | 59.29 | 23.73 | 8.75 | 4.17 | 1 |
| South Africa | 2 | 57.64 | 21.28 | 7.11 | 0.78 | 1 |
| Turkey | 2 | 45.56 | 24.71 | 12.36 | 8.51 | 1 |
| Panel Tests | | $r=0$ | $r=1$ | $r=2$ | $r=3$ | |
| C11 | | 2.41 | 4.39 | -0.18 | -1.56 | 2*** |
| C18 | | 5.66 | 5.67 | 0.62 | -2.68 | 2*** |
| C25 | | 6.58 | 6.28 | -0.15 | -1.30 | 2*** |
| Pedroni Tests | | panel-$\rho$ | panel-$t$ | group-$\rho$ | group-$t$ | |
| C11 | | 1.99 | -0.60 | 3.26 | -0.42 | |
| C18 | | 2.08 | -3.28*** | 3.73 | -2.96*** | |
| C25 | | 3.29 | -0.60 | 4.18 | -4.87*** | |

*Notes: ***, **, * Reject the null hypothesis at the 1%, 5%, 10% level.*

The average of (6.7) over time is denoted by

$$\frac{1}{T}\sum_{t=1}^{T} y_{it} = \delta_{0i} + \delta_{1i}\left(\frac{T+1}{2}\right) + \frac{1}{T}\sum_{t=1}^{T} x'_{it}\beta + \frac{1}{T}\sum_{t=1}^{T} z'_{it}\lambda_i + \frac{1}{T}\sum_{t=1}^{T} u_{it}. \quad (6.8)$$

To eliminate the individual-specific intercepts, (6.8) is subtracted from (6.7) which yields

$$\widetilde{y}_{it} = \delta_{1i}\widetilde{t} + \widetilde{x}'_{it}\beta + \widetilde{z}'_{it}\lambda_i + \widetilde{u}_{it}. \quad (6.9)$$

Note that $\widetilde{y}_{it} = y_{it} - \frac{1}{T}\sum_{t=1}^{T} y_{it}$, $\widetilde{x}_{it} = x_{it} - \frac{1}{T}\sum_{t=1}^{T} x_{it}$, $\widetilde{z}_{it} = z_{it} - \frac{1}{T}\sum_{t=1}^{T} z_{it}$, $\widetilde{u}_{it} = u_{it} - \frac{1}{T}\sum_{t=1}^{T} u_{it}$ and $\widetilde{t} = t - \frac{T-1}{2}$. To derive the panel DOLS estimator of $\theta = (\beta', \delta'_N, \lambda'_1, \ldots, \lambda'_N)'$ with $\delta_N = (\delta_{11}, \delta_{12}, \ldots, \delta_{1N})'$, the following matrices are defined.

$$\begin{aligned}
\widetilde{q}'_{1t} &= \begin{pmatrix} \widetilde{x}'_{1t} & \widetilde{t} & 0 & \ldots & 0 & \widetilde{z}'_{1t} & 0' & \ldots & 0' \end{pmatrix} \\
\widetilde{q}'_{2t} &= \begin{pmatrix} \widetilde{x}'_{2t} & 0 & \widetilde{t} & \ldots & 0 & 0' & \widetilde{z}'_{2t} & \ldots & 0' \end{pmatrix} \\
&\vdots \\
\widetilde{q}'_{Nt} &= \begin{pmatrix} \widetilde{x}'_{Nt} & 0 & \ldots & 0 & \widetilde{t} & 0' & \ldots & 0' & \widetilde{z}'_{Nt} \end{pmatrix}
\end{aligned} \quad (6.10)$$

According to this definition the panel DOLS estimator for $\theta$ is

$$\widehat{\theta}_{NT} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T} \widetilde{q}_{it}\widetilde{q}'_{it}\right)^{-1} \left(\sum_{i=1}^{N}\sum_{t=1}^{T} \widetilde{q}_{it}\widetilde{y}_{it}\right), \quad (6.11)$$

which has a limiting Gaussian distribution as $T \to \infty$ followed by $N \to \infty$.

As outlined in Section 6.1, according to economic theory the income elasticity should have a positive significant effect on the real money demand, whereas the semi-elasticities of the interest rate and the inflation rate should be negative. In Table 6.14 the estimates of the unknown parameters from Equation (6.3) are depicted for different number of cross-sections and time observations with different opportunity cost variables. For panel datasets with $T < 50$, the number of leads and lags is set to two, i.e. $p_i = 2$ for all $i = 1, \ldots, N$, and for panel datasets with $T < 30$, $p_i = 1$, for all $i = 1, \ldots, N$.

**Table 6.14:** Panel DOLS estimates of long-run money demand with an intercept and a linear time trend in the regression equation.

| Dataset | | OC | $\widehat{\beta}_1$(s.e) | $\widehat{\beta}_2$(s.e) | $\widehat{\beta}_3$(s.e) |
|---|---|---|---|---|---|
| 88/Q1-97/Q4 | C11 | $R^l$ | **0.86 (0.14)** | **-0.02 (0.01)** | |
| | | $R^s$ | **0.82 (0.18)** | **-0.02 (0.01)** | |
| | | $R^l, \pi$ | 0.74 (0.39) | **-0.02 (0.01)** | 1.10 (1.23) |
| | | $R^s, \pi$ | **0.76 (0.24)** | **-0.02 (0.01)** | 0.55 (0.88) |
| | C17 | $R^l$ | **1.17 (0.21)** | -0.02 (0.05) | |
| | | $R^s$ | 1.10 (0.81) | -0.02 (0.01) | |
| | | $R^l, \pi$ | **1.09 (0.39)** | **-0.02 (0.08)** | -0.04 (0.56) |
| | | $R^s, \pi$ | **1.14 (0.23)** | **-0.02 (0.01)** | -0.19 (0.63) |
| 90/Q1-97/Q2 | C18 | $R^l$ | **1.02 (0.17)** | -0.03 (0.02) | |
| | | $R^s$ | **0.84 (0.40)** | -0.00 (0.01) | |
| | | $R^l, \pi$ | **1.01 (0.38)** | -0.03 (0.02) | 0.73 (1.18) |
| | | $R^s, \pi$ | **0.87 (0.38)** | -0.00 (0.01) | -0.09 (1.34) |
| 93/Q1-97/Q2 | C11 | $R^s$ | -1.00 (1.90) | -0.02 (0.03) | |
| | | $R^s, \pi$ | -0.81 (2.42) | -0.02 (0.02) | 0.86 (2.60) |
| | C18 | $R^s$ | 0.48 (0.47) | -0.02 (0.02) | |
| | | $R^s, \pi$ | 0.47 (0.39) | -0.02 (0.02) | 0.41 (1.06) |
| | C25 | $R^s$ | **1.07 (0.49)** | 0.00 (0.00) | |
| | | $R^s, \pi$ | **0.87 (0.30)** | -0.00 (0.00) | 0.31 (0.67) |

*Notes: The figures in parentheses are the asymptotic standard errors. Bolded estimates are significant at least at the 5% significance level.*

The estimation results in Table 6.14 reveal that the income elasticity is mainly positive and significant if the time dimension is large enough or $N$ is high. The semi-elasticity of the interest rate is negative when the number of observations is $T = 40$. The semi-elasticity with respect to the inflation rate is not significant at all. This might be due to the fact that the inflation rates are $I(0)$ according to the panel unit root test results. This affects the estimates because the panel DOLS estimator is based on the assumption that the right hand side regressors are $I(1)$. This may be also the reason for the insignificant estimates of the semi-elasticities with respect to interest rate, especially for the sample 1993/Q1-1997/Q2. Hence, it can be concluded that the long-run money demand relation exists. Moreover, the point estimates for the income elasticity are around 1 and the estimates for the interest semi-elasticity are around -0.02 which coincides with the results of Mark and Sul (2003).

## 6.4 Conclusions

Throughout this chapter the existence of a long-run M1 demand relation was evaluated using panel unit root and cointegration tests. Additionally, the relation was estimated with the panel DOLS method of Mark and Sul (2003). The logarithm of the real income was taken as scale variable. The opportunity cost variables were chosen among the short-term and long-term interest rates and the annualized inflation rate.

To conclude on the robustness of the systems panel cointegration test statistic introduced in Chapter 5, the test was applied to different combinations of $T$ and $N$, dependent on the available data.

The order of the integratedness of individual variables was detected, with the help of the LLC and IPS panel unit root tests. It was found out that except the inflation rate, the variables are integrated of order one.

The panel SL cointegration test found evidence for the existence of one cointegrating relation and even more than one cointegrating relation using the inflation rate as a second opportunity cost variable, allowing this due to stationarity of inflation rates. The parametric tests of Pedroni (1999) detected the existence of only one cointegrating relation, but just for large $N$.

Finally, by the panel DOLS estimator of Mark and Sul (2003), for panel datasets with large $T$, a positive income elasticity and a negative semi-elasticity with respect to interest rates were discovered.

# Appendix A

# Index for Notation

### Abbreviations

| | |
|---|---|
| ADF | Augmented Dickey-Fuller |
| AR($p$) | autoregressive process of order $p$ |
| CPI | consumer price index |
| CUSUM | cumulative sum |
| DF | Dickey-Fuller |
| DGP | data generating process |
| DOLS | dynamic ordinary least squares |
| ECB | European Central Bank |
| EU | European Union |
| EMU | Economic and Monetary Union |
| FMOLS | Fully-modified ordinary least squares |
| FOLS | feasible ordinary least squares |
| GCC | Gulf Cooperation Council |
| GDP | gross domestic product |
| GLS | generalized least squares |
| GMM | generalized method of moments |
| IMF | International Monetary Fund |
| IPS | Im, Pesaran & Shin |
| IV | instrumental variable |
| LLC | Levin, Lin & Chu |
| LLL | Larson, Lyhagen & Löthgren |
| LM | Lagrange multiplier |
| LR | likelihood ratio |
| LSDV | least squares dummy variable |

| | |
|---|---|
| LBUI | locally best unbiased invariant |
| MA($q$) | moving average process of order $q$ |
| OECD | Organisation for Economic Co-operation and Development |
| OLS | ordinary least squares |
| RR | reduced rank |
| SL | Saikkonen & Lütkepohl |
| UK | United Kingdom |
| US | United States |
| VAR | vector autoregressive |
| VAR($p$) | vector autoregressive of order $p$ |
| VEC | vector error correction |
| VECM | vector error correction model |

**Symbols and Operators**

| | |
|---|---|
| $\sim$ | is distributed as |
| $\rightarrow$ | converges to |
| $\stackrel{w}{\rightarrow}$ | converges weakly to |
| $\infty$ | infinity |
| $\in$ | is an element of |
| $\varnothing$ | empty set |
| $E(y)$ | expectation of $y$ |
| $Var(y)$ | variance of $y$ |
| $Cov(y)$ | covariance matrix of y |
| $Cov(x, y)$ | covariance between x and y |
| $\Delta$ | differencing operator: $\Delta y_{it} = y_{it} - y_{i,t-1}$ |
| $\sum$ | summation sign |
| $\int_a^b$ | definite integral from $a$ to $b$ |
| $\ell(\cdot)$ | likelihood function |
| $\ln \ell(\cdot)$ | log-likelihood function |
| ln | natural logarithm |
| lim | limit |
| $\|y\|$ | absolute value or modulus of $y$ |
| $\|\|y\|\|$ | norm of $y$ |
| $[y]$ | integer part of $y$ |
| $I(d)$ | integrated process of order $d$, which is stationary after differencing $d$ times |

| | |
|---|---|
| $L$ | lag operator: $Ly_{it} = y_{i,t-1}$ |
| min | minimum of |
| max | maximum of |
| sup | supremum, least upper bound |
| arg min $(X)$ | argument that minimizes expression X |
| $(m \times n)$ | dimension of a matrix, $m$ is the number of rows and $n$ is the number of columns |
| $I_T$ | identity matrix of dimension $(T \times T)$ |
| $\mathbb{1}_T$ | $T$-dimensional column vector of ones |
| $J_T$ | $(T \times T)$ matrix of ones |
| $A'$ | transpose of a matrix $A$ |
| $A^{-1}$ | inverse of a matrix $A$ |
| diag$(a_1, \ldots, a_n)$ | diagonal matrix with diagonal elements $a_1, \ldots, a_n$ |
| tr(A) | trace of a square matrix A |
| det(A) | determinant of a square matrix A |
| vec | column stacking operator |
| $\mathrm{A}_\perp$ | orthogonal complement of an $(m \times n)$ matrix A with $\mathrm{A}_\perp$ $(m \times (m-n))$, such that both, A and $\mathrm{A}_\perp$, have full column rank, and $\mathrm{A}'\mathrm{A}_\perp = 0$ |
| rank(A) | rank of a $(m \times n)$ matrix A |
| $\otimes$ | kronecker product |
| *df* | degrees of freedom |
| *i.i.d.* | independently, identically distributed |
| $N_m(\mu, \Sigma)$ | $m$-variate normal distribution with mean $\mu$ and variance (covariance matrix) $\Sigma$ |
| $W_m(n, \Sigma)$ | $(m \times m)$ matrix variate Wishart distribution with $n$ degrees of freedom and covariance matrix $\Sigma$ |
| $\chi^2_{(m)}$ | chi-square distribution with m degrees of freedom |
| $t_{(m)}$ | $t$-distribution with $m$ degrees of freedom |
| $U(a, b)$ | uniform distribution on the interval $[a, b]$ |
| $T$ | time series lenght |
| $N$ | cross-section length |
| $O(n^k)$ | at most of order $n^k$ |

| | |
|---|---|
| $o(n^k)$ | of smaller order than $n^k$ |
| $O_p(n^k)$ | at most of order $n^k$ in probability |
| $W(s)$ | multivariate standard Brownian motion with identity covariance matrix |
| $H_0$ | null hypothesis |
| $H_1$ | alternative hypothesis |
| $r$ | cointegrating rank |

# Appendix B

# Data Sources

## B.1 Fisher Hypothesis

According to the availability of the data monthly panel datasets with two different time spans, i.e from June 1989 to December 1998 ($T = 115$), and from January 1991 to December 2002 ($T = 145$), were used for the following variables from the given sources.

- **Price index:** Consumer Price Index (CPI) for All Items (Base 1995= 100) from *Organisation for Economic Co-operation and Development* (OECD). For Ireland Wholesale Price Index is used. The variable $p_{it}$ represents the logarithm of the price index for cross-section $i$ at time $t$.

- **Inflation Rate:** Three-month inflation rate calculated by $\pi_{it} = p_{it} - p_{i,t-3}$. The variable $\pi_{it}$ denotes the inflation rate for cross-section $i$ at time $t$.

- **Interest Rate:** Three-month interest rates from *OECD*. The variable $n_{it}$ is the short-term interest rate for cross-section $i$ at time $t$.

The first panel dataset (dataset A) consists of the following countries: Austria, Belgium, Canada, Denmark, France, Finland, Germany, Iceland, Ireland, Italy, Japan, Mexico, Netherlands, Norway, Portugal, Spain, Sweden, UK and US ($N = 19$).

The second panel dataset (dataset B) consists of the following countries: Canada, Denmark, Hungary, Iceland, Japan, Korea, Mexico, Norway, Sweden, UK and US ($N = 11$).

## B.2 Money Demand Function

Quarterly panel dataset with three different time spans, i.e from the first quarter of 1988 to the fourth quarter of 1997 ($T = 40$), from the first quarter of 1990 to the second quarter of 1997 ($T = 30$) and from the first quarter of 1993 to the second quarter of 1997 ($T = 18$), were used for the following variables from the given sources.

- **Price index:** Consumer Price Index (CPI) for All Items (Base 2000= 100) is the price level. The variable $p_{it}$ denotes the logarithm of the price index for cross-section $i$ at time $t$.

- **Money level:** M1 is the nominal money variable. The variable $m_{it}$ is the logarithm of the nominal money level for cross-section $i$ at time $t$.

- **Income:** Gross domestic product (GDP). The variable $g_{it}$ is the logarithm of the real income for cross-section $i$ at time $t$.

- **Long-term interest rate:** Average government bond yield. The variable $R^l_{it}$ represents the long-term interest rate for cross-section $i$ at time $t$.

- **Short-term interest rate:** Average deposit rate. The variable $R^s_{it}$ denotes the short-term interest rate for cross-section $i$ at time $t$.

- **Annual inflation rate:** The variable $\pi_{it}$ is the annualized inflation rate for cross-section $i$ at time $t$, i.e. $\pi_{it} = 4(p_{it} - p_{i,t-1})$.

The countries included in the panel datasets are coded as follows:

- C11: Austria, Belgium, Finland, France, Germany, Italy, Netherlands, Norway, Portugal, Spain, Switzerland,

- C17 : C11 countries and Australia, Canada, Japan, Korea, New Zealand, US,

- C18 : C17 countries and Denmark,

- C25 : C18 countries and Argentina, Brazil, Indonesia, Malaysia, Mexico, South Africa, Turkey.

Table B.1 presents the data sources of each variable for different cross-sections. (s.a.) refers to the variables gathered seasonally-adjusted from the data source. With the seasonal-adjustment method XCensus 11 the seasonally-unadjusted variables are also adjusted. The data are provided from two main sources, i.e. *International Monetary Fund* (IMF) and *OECD*.

**Table B.1:** Data sources for the money demand function.

| Country | M1 | GDP | $R^l$ | $R^s$ | CPI |
|---|---|---|---|---|---|
| Argentina | IMF (s.a.) | IMF | | IMF | IMF |
| Australia | IMF (s.a.) | IMF (s.a.) | OECD | IMF | IMF |
| Austria | IMF | IMF | IMF | IMF | IMF |
| Belgium | OECD | OECD | IMF | IMF | IMF |
| Brazil | IMF | IMF | | IMF | IMF |
| Canada | IMF (s.a.) | IMF(s.a.) | IMF | IMF | IMF |
| Denmark | OECD (s.a.) | OECD (s.a.) | IMF | | |
| Finland | OECD | OECD (s.a) | OECD | IMF | OECD |
| France | IMF (s.a.) | IMF (s.a.) | IMF | IMF | IMF |
| Germany | IMF (s.a.) | IMF (s.a.) | IMF | IMF | OECD |
| Indonesia | OECD (s.a.) | OECD(s.a.) | | IMF | OECD |
| Italy | OECD (s.a) | OECD (s.a) | IMF | IMF | OECD |
| Japan | IMF (s.a) | IMF (s.a.) | IMF | IMF | IMF |
| Korea | OECD (s.a.) | OECD (s.a) | IMF | IMF | OECD |
| Malaysia | IMF | IMF | | IMF | IMF |
| Mexico | OECD | OECD (s.a.) | | IMF | OECD |
| Netherlands | OECD | OECD (s.a.) | IMF | IMF | OECD |
| New Zealand | OECD (s.a) | IMF (s.a) | IMF | IMF | IMF |
| Norway | OECD | OECD (s.a.) | IMF | IMF | OECD |
| Portugal | IMF | OECD | IMF | IMF | OECD, IMF |
| South Africa | OECD (s.a.) | OECD | | IMF | OECD |
| Spain | IMF | IMF (s.a.) | IMF | IMF | IMF |
| Switzerland | OECD | OECD (s.a.) | IMF | IMF | OECD |
| Turkey | OECD (s.a.) | OECD (s.a.) | | IMF | OECD |
| US | IMF (s.a.) | IMF (s.a.) | IMF | IMF | IMF |

# Bibliography


S. K. Ahn. Some tests for unit root in autoregressive-integrated moving average models with deterministic trends. *Biometrica*, 80:855–868, 1993.

S. K. Ahn and G. C. Reinsel. Estimation for partially nonstationary multivariate autoregressive models. *Journal of American Statistical Association*, 85:813–823, 1990.

C. Amsler and J. Lee. An LM test for a unit root in the presence of a structural break. *Econometric Theory*, 11:359–368, 1995.

R. Anderson, H. Qian, and R. Rasche. Analysis of panel vector error correction models using maximum likelihood, bootstrap, and canonical-correlation estimators. Working Paper Series 2006-050A, Federal Reserve Bank of St. Louis, August 2006.

F. J. Atkins. Cointegration, error correction and Fisher effect. *Applied Economics*, 21:1611–1620, 1989.

J. Bai and S. Ng. A PANIC attack on unit roots and cointegration. *Econometrica*, 72:1127–1177, 2004.

J. Bai and P. Perron. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18:1–22, 2003.

J. Bai and P. Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66:47–78, 1998.

A. Banerjee and J. L. Carrion-i- Silvestre. Cointegration in panel data with breaks and cross-section dependence. Working Paper Series 591, European Central Bank, February 2006.

A. Banerjee, J. J. Dolado, and R. Mestre. Error-correction mechanism tests for cointegration in a single-equation framework. *Journal of Time Series Analysis*, 19:267–283, 1998.

A. Banerjee, M. Marcellino, and C. Osbat. Some cautions on the use of panel methods for integrated series of macroeconomic data. *Econometrics Journal*, 7:322–340, 2004.

R. Bewley and M. Yang. Tests for cointegration based on canonical correlation analysis. *Journal of the American Statistical Association*, 90:990–996, 1995.

H. P. Boswijk. Testing identifiability of cointegrated systems. *Journal of Business and Economic Statistics*, 14:153–160, 1996.

G. E. P. Box and G. C. Tiao. A canonical analysis of multiple time series. *Biometrika*, 64:355–365, 1977.

J. Breitung. Nonparametric tests for unit root and cointegration. *Journal of Econometrics*, 108:343–363, June 2002.

J. Breitung. A parametric approach to the estimation of cointegration vectors in panel data. *Econometric Reviews*, 24:1–20, 2005.

J. Breitung and W. Meyer. Testing for unit roots in panel data: Are wages on different bargaining levels cointegrated? *Applied Economics*, 61:353–361, 1994.

C. Carrera. Long-run money demand in Latin-American countries: A nonstationary panel data approach. Technical report, Central Reserve Bank of Peru, 2006.

Y. Chang. Nonlinear IV unit root tests in panels with cross-sectional dependency. *Journal of Econometrics*, 110:261–292, 2002.

Y. Chang. Bootstrap unit root tests in panels with cross-sectional dependency. *Journal of Econometrics*, 120:263–293, 2004.

Y. Chang and J. Y. Park. A sieve bootstrap for the test of a unit root. *Journal of Time Series Analysis*, 24:379–400, 2003.

Y. Chang, J. Y. Park, and K. Song. Bootstrapping cointegrating regressions. *Journal of Econometrics*, 133:703–739, 2006.

S. S. Cheng. *Partial Difference Equations*. London and New York: Taylor and Francis, 2003.

I. Choi. Unit root tests for panel data. *Journal of International Money and Finance*, 20:249–272, 2001.

I. Choi. Improving the empirical size of the KPSS test of stationarity. Hong Kong University of Science and Technology, Department of Economics, January 2004.

D. Conniffe and J. E. Spencer. When moments of ratios are ratios of moments. *The Statistician*, 50:161–168, 2001.

M. Costantini and L. Gutierrez. Simple panel unit root tests to detect changes in persistence. *Economics Letters*, 96:363–368, 2007.

W. J. Crowder. Panel estimates of the Fisher effect. Working Paper Series 03-003, University of Texas at Airlington, Department of Economics, 2003.

W. J. Crowder. The long-run Fisher relation in Canada. *Canadian Journal of Economics*, 30:1124–1142, 1997.

W. J. Crowder and D. L. Hoffman. The long-run relationship between nominal interest rates and inflation: The Fisher equation revisited. *Journal of Money, Credit and Banking*, 28:102–118, 1996.

F. Di Iorio and S. Fachin. Cointegration testing in dependent panels with breaks. MPRA Paper 3139, University Library of Munich, Germany, 2007.

C. Dreger, H.-E. Reimers, and B. Roffia. Long-run money demand in the new EU member states with exchange rate effects. *Eastern European Economics*, 45:75–94, 2007.

S. Dutt and D. Gosh. The Fisher hypothesis: Examining the Canadian experience. *Applied Economics*, 27:1025–1030, 1995.

R. F. Engle and C. W. J. Granger. Cointegration and error correction: Representation, estimation and testing. *Econometrica*, 55:251–276, 1987.

M. D. Evans and K. K. Lewis. Do expected shifts in inflation affect estimates of the long-run Fisher relation? *Journal of Finance*, 50:225–253, 1995.

S. Fachin. Long-run trends in internal migrations in Italy: A study in panel cointegration with dependent units. *Journal of Applied Econometrics*, 22: 401–428, 2007.

J. Fidrmuc. Money demand and disinflation in selected CEECs during the accession to the EU. *forthcoming in Applied Economics*, 2008.

I. Fisher. *The Theory of Interest.* New York: MacMillan, 1930.

R. A. Fisher. *Statistical Methods For Research Workers*. Oliver and Boyn, Edinburgh, 1932.

C. Gengenbach, F. C. Palm, and J.-P. Urbain. Cointegration testing in panels with common factors. *Oxford Bulletin of Economics and Statistics*, 68:683–719, 2006.

N. P. v. Giersbergen. Bootstrapping the trace statistic in VAR models: Monte Carlo results and applications. *Oxford Bulletin of Economics and Statistics*, 58:631–656, 1996.

J. Gonzalo. Five alternative methods of estimating long-run equilibrium relationships. *Journal of Econometrics*, 60:203–233, 1994.

J. Gonzalo and C. W. J. Granger. Estimation of common long-memory components in cointegrated systems. *Journal of Business and Economic Statistics*, 13:27–36, 1995.

P. Graczyk, G. Letac, and H. Massam. The hyperoctahedral group, symmetric group representations and the moments of the real Wishart distribution. *Journal of Theoretical Probability*, 18:1–42, 2005.

A. M. Gregory and B. E. Hansen. Residual-based tests for cointegration in models with regime shifts. *Journal of Econometrics*, 70:99–126, 1996a.

A. M. Gregory and B. E. Hansen. Tests for cointegration in models with regime and trend shifts. *Oxford Bulletin of Economics and Statistics*, 58: 555–560, 1996b.

A. M. Gregory, A. A. Haug, and N. Lomuto. Mixed signals among tests for cointegration. *Journal of Applied Econometrics*, 19:89–98, 2004.

J. J. J. Groen and F. Kleibergen. Likelihood-based cointegration analysis in panels of vector error correction models. *Journal of Business and Economic Statistics*, 21:295–318, 2003.

L. Gutierrez. On the power of panel cointegration tests: A Monte Carlo comparison. *Economics Letters*, 80:105–111, 2003.

L. Gutierrez. Simple tests for cointegration in panels with structural breaks. *forthcoming in Applied Economics Letters*, 2008.

C. Hanck. The error-in-rejection probability of meta-analytic panel tests. *forthcoming in Economics Letters*, 2008.

C. Hanck. Mixed signals among panel cointegration tests. Universität Dortmund, SFB 475 Technical Report 45/06, Ruhr Graduate School in Economics, November 2006a.

C. Hanck. Cross-sectional correlation robust tests for panel cointegration. Universität Dortmund, SFB 475 Technical Report 44/06, Ruhr Graduate School in Economics, November 2006b.

C. Hanck. A meta analytic approach to testing for panel cointegration. Universität Dortmund, SFB 475 Technical Report 02/07, Ruhr Graduate School in Economics, January 2007.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982.

N. Harb. Money demand function: A heterogeneous panel application. *Applied Economics Letters*, 11:551–555, 2004.

D. Harris and B. Inder. A test of the null hypothesis of cointegration. In C. P. Hargreaves, editor, *Nonstationary Time Series Analysis and Cointegration*. Oxford University Press, New York, 1994.

A. A. Haug. Tests for cointegration: A Monte Carlo comparison. *Journal of Econometrics*, 71:89–115, 1996.

J. Hlouskova and M. Wagner. The performance of panel unit root and stationarity tests: Results from a large scale simulation study. Economics Working Papers 2005/05, European University Institute, 2005.

K. Hubrich, H. Lütkepohl, and P. Saikkonen. A review of systems of cointegration tests. *Econometric Reviews*, 20:247–318, 2001.

K. S. Im, H. M. Pesaran, and Y. Shin. Testing for unit roots in heterogeneous panels. Technical report, Trinity College Cambridge, 1997.

K. S. Im, H. M. Pesaran, and Y. Shin. Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115:53–74, 2003.

K. S. Im, J. Lee, and M. Tieslau. Panel LM unit-root tests with level shifts. *Oxford Bulletin of Economics and Statistics*, 67:393–419, 2005.

S. Johansen. Statistical analysis of cointegrating vectors. *Econometric Reviews*, 12:231–254, 1988.

S. Johansen. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59:1551–1580, 1991.

S. Johansen. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press: Oxford, 1995.

N. L. Johnson and S. Kotz. *Continuous Univariate Distributions-1*. New York: Wiley, 1970a.

N. L. Johnson and S. Kotz. *Continuous Univariate Distributions-2*. New York: Wiley, 1970b.

M. C. Jones. Expressions for inverse moments of positive quadratic forms in normal variables. *The Australian Journal of Statistics*, 28:242–250, 1986.

M. C. Jones. On the moments of ratios of quadratic forms in normal variables. *Statistics and Probability Letters*, 6:129–136, 1987.

K. Jönsson. Using panel data to increase the power of modified unit root tests in the presence of structural breaks. *Applied Mathematics and Computation*, 171:832–842, 2005.

G. G. Judge, R. C. Hill, W. E. Griffiths, H. Lütkepohl, and T.-C. Lee. *Introduction to the Theory and Practice of Econometrics*. New York: Wiley, 1988.

C. Kao. Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics*, 90:1–44, 1999.

C. Kao and M.-H. Chiang. On the estimation and inference of a cointegrated regression in panel data. *Advances in Econometrics: Nonstationary Panels, Panel Cointegration and Dynamic Panels*, 15:179–222, 2000.

D. D. Karaman Örsal. Comparison of panel cointegration tests. *Economics Bulletin*, 3:1–20, 2008.

S. Karlsson and M. Löthgren. On the power and interpretation of panel unit root tests. *Economics Letters*, 66:249–255, 2000.

R. Larsson. The order of the asymptotic error term for moments of the log likelihood ratio test for cointegration. unpublished manuscript, Department of Statistics, Stockholm University, 1999.

R. Larsson and J. Lyhagen. Likelihood-based inference in multivariate panel cointegration models. Working Paper Series in Economics and Finance 331, Stockholm School of Economics, November 1999.

R. Larsson, J. Lyhagen, and M. Löthgren. Likelihood-based cointegration tests in heterogeneous panels. *Econometrics Journal*, 4:109–142, 2001.

G. Letac and H. Massam. All the moments and the inverse moments of the Wishart distribution. *Seminarberichte aus dem Fachbereich Mathematik*, 68:309–339, 1999.

A. Levin and C. F. Lin. Unit root tests in panel data: New results. Discussion Paper 93-56, University of California, San Diego, 1993.

A. Levin, C. F. Lin, and C.-S. James Chu. Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics*, 108: 1–24, 2002.

H. Lütkepohl. *New Introduction To Multiple Time Series Analysis*. Berlin: Springer, 2005.

H. Lütkepohl and M. Krätzig, editors. *Applied Time Series Econometrics*. Cambridge University Press, Cambridge, 2004.

H. Lütkepohl and P. Saikkonen. Testing for the cointegrating rank of a VAR process with a time trend. *Journal of Econometrics*, 95:177–198, 2000.

R. MacDonald and P. D. Murphy. Testing for the long run relationship between nominal interest rates and inflation using cointegration techniques. *Applied Economics*, 21:439–47, April 1989.

J. G. MacKinnon. Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics*, 11:601–618, 1996.

G. S. Maddala and S. Wu. A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics*, 61:631–652, 1999.

N. C. Mark and D. Sul. Cointegration vector estimation by panel DOLS and long-run money demand. *Oxford Bulletin of Economics and Statistics*, 65: 655–680, 2003.

A. M. Mathai and S. B. Provost. *Quadratic Forms in Random Variables: Theory and Applications*. New York: Marcel Dekker, 1992.

S. McCoskey and C. Kao. A residual-based test of the null of cointegration in panel data. *Econometric Reviews*, 17:57–84, 1998.

S. McCoskey and C. Kao. Comparing panel data cointegration tests with an application to the "Twin Deficits" problem. *Working Paper, Center for Policy Research*, 1999.

F. S. Mishkin. Is the Fisher effect for real?: A reexamination of the relationship between inflation and interest rates. *Journal of Monetary Economics*, 30:195–215, November 1992.

R. J. Muirhead. *Aspects of Multivariate Statistical Theory.* New York: Wiley, 1982.

M. Neumann. Inverses of Perron complements of inverse M-matrices. *Linear Algebra and its Applications*, 313:163–171, 2000.

W. K. Newey and K. D. West. A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 50: 703–708, 1987.

W. K. Newey and K. D. West. Autocovariance lag selection in covariance matrix estimation. *Review of Economic Studies*, 61:631–653, 1994.

J. Nyblom and A. Harvey. Testing for common stochastic trends. *Econometric Theory*, 16:176–199, 2000.

E. Paparoditis and D. Politis. The Continuous-Path Block Bootstrap. In M. Puri, editor, *Asymptotics in Statistics and Probability. Papers in honor of George Roussas.* VSP Publications, Zeist, 2001.

P. Pedroni. Fully-modified OLS for heterogeneous cointegrated panels and the case of purchasing power parity. *Advances in Econometrics*, 15:93–130, 2000.

P. Pedroni. Panel cointegration; asymptotic and finite sample properties of pooled time series tests, with an application to the PPP hypothesis. *Econometric Theory*, 20:575–625, 2004.

P. Pedroni. Panel cointegration; asymptotic and finite sample properties of pooled time series tests, with an application to the PPP hypothesis. Working Paper Series in Economics 95-013, Indiana University, 1995.

P. Pedroni. Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics*, 61: 653–670, 1999.

P. Perron. Testing for a random walk: A simulation experiment of power when the sampling interval is varied. In B. Jaj, editor, *Advances in Econometrics and Modeling.* Kluwer Academic Publishers, 1989.

P. Perron. Test consistency with varying sampling frequency. *Econometric Theory*, 7:341–368, 1991.

P. C. B. Phillips and B. E. Hansen. Statistical inference in instrumental variables regression with $I(1)$ process. *Review of Economics Studies*, 57: 99–125, January 1990.

P. C. B. Phillips and M. Loretan. Estimating long-run economic equilibria. *Review of Economic Studies*, 58:407–436, 1991.

P. C. B. Phillips and H. R. Moon. Linear regression limit theory for nonstationary panel data. *Econometrica*, 67:1057–1111, 1999.

P. C. B. Phillips and S. Ouliaris. Asymptotic properties of residual based tests for cointegration. *Econometrica*, 58:165–193, 1990.

P. C. B. Phillips and P. Perron. Testing for a unit root in time series regressions. *Biometrika*, 75:335–346, 1988.

P. C. B. Phillips and V. Solo. Asymptotics of linear processes. *The Annals of Statistics*, 20:971–1001, 1992.

R. G. Pierce and A. J. Snell. Temporal aggregation and the power of tests for a unit root. *Journal of Econometrics*, 65:333–345, 1995.

D. Quah. Exploiting cross section variation for unit root inference in dynamic data. *Economics Letters*, 44:9–19, 1994.

A. K. Rose. Is the real interest rate stable? *Journal of Finance*, 43:1095–1112, December 1988.

P. Saikkonen. Asymptotic efficient estimation of cointegration regressions. *Econometric Theory*, 7:1–21, 1991.

P. Saikkonen. Testing the normalization and overidentification of cointegrating vectors in vector autoregressive processes. *Econometric Reviews*, 18: 235–257, 1999.

P. Saikkonen and H. Lütkepohl. Local power of likelihood ratio tests for the cointegrating rank of a VAR process. *Econometric Theory*, 15:50–78, 1999.

P. Saikkonen and H. Lütkepohl. Trend adjustment prior to testing for the cointegrating rank of a vector autoregressive process. *Journal of Time Series Analysis*, 21:435–456, 2000a.

P. Saikkonen and H. Lütkepohl. Testing for the cointegrating rank of a VAR process with structural shifts. *Journal of Business and Economic Statistics*, 18:451–464, 2000b.

P. Schmidt and P. C. B. Phillips. LM tests for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics*, 54: 257–287, 1992.

R. J. Serfling. *Approximation Theorems of Mathematical Statistics*. New York: Wiley, 1980.

R. Shiller and P. Perron. Testing the random walk hypothesis: Power versus frequency of observation. *Economic Letters*, 18:381–386, 1985.

Y. Shin. A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory*, 10:91–115, 1994.

S. S. Sriram. A survey of recent empirical money demand studies. *IMF Staff Papers*, 47:334–365, 2001.

J. H. Stock and M. W. Watson. Testing for common trends. *Journal of American Statistical Association*, 83:1097–1107, 1988.

J. H. Stock and M. W. Watson. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica*, 61:783–820, July 1993.

P. S. Tam. Testing for no cointegration with structural breaks in panel data. University of Macau, China, Faculty of Business Administration, 2007.

H. Y. Toda. Finite sample properties of likelihood ratio tests for cointegrating ranks when linear trends are present. *Review of Economics & Statistics*, 66:66–79, 1994.

H. Y. Toda. Finite sample performance of likelihood ratio tests for cointegrating ranks in vector autoregressions. *Econometric Theory*, 11:1015–1032, 1995.

C. Trenkler. Bootstrapping systems cointegration tests with a prior adjustment for deterministic terms. *forthcoming in Econometric Theory*, 2008.

C. Trenkler, P. Saikkonen, and H. Lütkepohl. Testing for the cointegrating rank of a VAR process with level shift and trend break. *Journal of Time Series Analysis*, 29:331–358, 2008.

M. Wagner and J. Hlouskova. The performance of panel cointegration methods: Results from a large scale simulation study. Economics Series 210, Institute for Advanced Studies, Vienna, May 2007.

J. Westerlund. New simple tests for panel cointegration. *Econometric Reviews*, 24:297–316, 2005a.

J. Westerlund. A panel CUSUM test of the null of cointegration. *Oxford Bulletin of Economics and Statistics*, 67:231–262, 2005b.

J. Westerlund. Reducing the size distortion of the panel LM test for cointegration. *Economics Letters*, 90:384–389, 2006a.

J. Westerlund. Testing for panel cointegration with multiple structural breaks. *Oxford Bulletin of Economics and Statistics*, 68:101–132, 2006b.

J. Westerlund. Testing for panel cointegration with a level break. *Economics Letters*, 91:27–33, 2006c.

J. Westerlund. Testing for error correction in panel data. *Oxford Bulletin of Economics and Statistics*, 69:709–748, 2007.

J. Westerlund. Panel cointegration tests of the Fisher Hypothesis. *Journal of Applied Econometrics*, 23:193–233, 2008.

J. Westerlund and S. A. Basher. Mixed signals among tests for panel cointegration. *forthcoming in Economic Modelling*, 2008.

J. Westerlund and D. L. Edgerton. A panel bootstrap cointegration test. *Economics Letters*, 97:185–190, 2007a.

J. Westerlund and D. L. Edgerton. Simple tests for cointegration in dependent panels with structural breaks. Lund University, Department of Economics, January 2007b.

Z. Xiao. A residual based test for the null hypothesis of cointegration. *Economics Letters*, 64:133–141, August 1999.

Z. Xiao and P. C. B. Phillips. A CUSUM test for cointegration using regression residuals. *Journal of Econometrics*, 108:43–61, May 2002.

W.-C. Yueh. Eigenvalues of several tridiagonal matrices. *Applied Mathematics E-Notes*, 5:66–74, 2005.

W.-C. Yueh. Explicit inverses of several tridiagonal matrices. *Applied Mathematics E-Notes*, 6:74–83, 2006.

# List of Figures

# List of Tables

# Selbständigkeitserklärung

Hiemit erkläre ich, dass ich die von mir vorgelegte Dissertation selbstständig angefertigt habe und alle benutzten Quellen und Hilfsmittel vollständig angegeben habe. Alle wörtlich oder inhaltlich übernommenen Stellen habe ich als solche gekennzeichnet.

Berlin, den 28.07.2008 Deniz Dilan Karaman Örsal